## (19) World Intellectual Property Organization International Bureau





#### (43) International Publication Date 25 April 2002 (25.04.2002)

#### PCT

### (10) International Publication Number WO 02/32943 A2

(51) International Patent Classification7: C07K 14/16 DEPARTMENT OF HEALTH AND HUMAN SER-VICES [US/US]; National Institutes of Health, Office of Technology Transfer, 6011 Executive Blvd., Suite 325, (21) International Application Number: PCT/US01/25721 Rockville, MD 20852-3804 (US). CHADRABARTI, Bimal, K. [IN/US]; 1147 McIntyre Drive, Ann Arbor, MI (22) International Filing Date: 14 August 2001 (14.08.2001) 48105 (US). (25) Filing Language: **English** 

- (26) Publication Language: English
- (30) Priority Data: 60/225,097 14 August 2000 (14.08.2000) US 60/252,115 14 November 2000 (14.11.2000) US 60/279,257 28 March 2001 (28.03.2001) US
- (71) Applicants (for all designated States except US): THE GOVERNMENT OF THE UNITED STATES OF AMERICA, as represented by THE SECRETARY,

- (72) Inventors; and
- (75) Inventors/Applicants (for US only): NABEL, Gary, J. [US/US]; 2520 30th Street, NW, Washington, DC 20008 (US). HUANG, Yue [CA/US]; 413 West Side Drive, #302, Gaithersburg, MD 20878 (US).
- (74) Agent: ALTMAN, Daniel, E.; Knobbe, Martens, Olson and Bear, LLP, 16th Floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA,

[Continued on next page]

(54) Title: MODIFICATIONS OF HIV ENV, GAG, AND POL ENHANCE IMMUNOGENICITY FOR GENETIC IMMUNIZA-TION



(57) Abstract: Modified IIIV Env., Gag. Pol., or Nef DNA with improved ability to elicit antibody and CTL responses to HIV antigens have been identified as prototype immunogens for the treatment and prevention of HIV infections.





core of HIV gp41

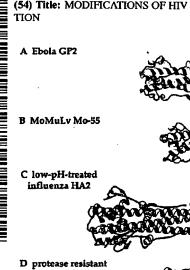


E SIV gp41 (NMR)



F SNARES







CH, CN, CO, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EC, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

#### Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

## Modifications of HIV Env, Gag, and Pol Enhance Immunogenicity for Genetic Immunization

5

### Field of the Invention

The present invention relates to the field of molecular biology. The present invention discloses modified HIV Env, Gag, Pol, and Nef proteins, related nucleotide sequences, and usage for genetic immunization.

#### **Background of the Invention**

10

15

20

25

30

35

Protective immunity against human immunodeficiency virus-1 (HIV-1) is likely to require recognition of linear and conformation epitopes from multiple HIV antigens. Whether these responses can be elicited more effectively by virion-like structures or fused CTL epitopes is unknown.

The immune response to HIV infection in long-term non-progressors (Cao Y et al., 1995, N Engl J Med, 332:201-8; Pantaleo G et al., 1995, N Engl J Med, 332:209-16) and HIV-exposed sex workers suggests that specific viral immunity may limit infection and the symptoms of disease. No single characteristic yet correlates with protective immunity, but studies in non-human of primates suggest that both humoral and cellular immunity are required for this response. Depletion of cytotoxic T cells (CTLs) in chronically-infected macaques enhances viremia (Jin X et al., 1999, J Exp Med, 189:991-8; Schmitz JE et al., 1999, Science, 283:857-60). In humans, higher CTL responses correlate with lower viral load and stabilization of clinical symptoms (Musey L et al., 1997, N Engl J Med, 337:1267-74; Ogg GS et al., 1998, Science, 279:2103-6). In animal models, passive transfer of neutralizing antibodies can also contribute to protection against virus challenge (Burton DR, Montefiori DC, 1997, AIDS, 11 Suppl A:S87-S98; Gauduin MC et al., 1997, Nat Med, 3:1389-93; Moore JP, Ho DD., 1995, AIDS, 9 Suppl A:S117-S136; Muster T et al., 1994, J Virol, 68:4031-4; Muster T et al., 1993, J Virol, 67:6642-7; Poignard P et al., 1996, Immunol Today, 17:239-46; Sattentau QJ, 1996, Curr Opin Immunol, 8:540-5; Shibata R et al., 1999, Nat Med, 5:204-10). Neutralizing antibody responses can also be developed in HIV-infected individuals (Burton DR, Montefiori DC., 1997, AIDS, 11 Suppl A:S87-S98; Poignard P et al., 1996, Immunol Today, 17:239-46; Sattentau QJ., 1996, Curr Opin Immunol, 8:540-5) and are associated with lower viral loads in long-term non-progressors (Montefiori DC et al., 1996, J Infect Dis, 173:60-7). Though this neutralizing antibody response is

uncommon, it is directed largely against the Env protein of the virus (Robey WG et al., 1986, Proc Natl Acad Sci U S A, 83:7023-7; Steimer KS et al., 1991, Science, 254:105-8).

In early human vaccine trials, gp120 protein immunogens have yielded disappointing results: vaccine-induced antibodies have not been broadly neutralizing and have sometimes enhanced infection in vitro (Bolognesi DP, Matthews TJ, 1998, Nature, 391:638-9; Connor RI et al., 1998, J Virol, 72:1552-76; Haynes BF, 1996, Lancet, 348:933-7; VanCott TC et al., 1995, J Immunol, 155:4100-10). Monomeric gp120 loses oligomer-dependent epitopes and does not include sequences in the ectodomain of the gp41 that become exposed during virus entry (Broder CC et al., 1994, Proc Natl Acad Sci U S A, 91:11699-703; Moore JP, 1995, Nature, 376:115; Muster T et al., 1993, J Virol, 67:6642-7). It is assumed that broadly neutralizing antibodies bind to native gp120/gp41 complex on the surface of the virus rather than soluble gp120 (Burton DR, Montefiori DC, 1997, AIDS, 11 Suppl A:S87-S98).

15

20

25

30

35

10

5

The development of a cytotoxic T lymphocyte (CTL) response to viruses is often crucial to the outcome of infections. Lysis of infected cells prior to the production of progeny virions may limit virus burst size (Yang, O et al., 1996, J. Virol., 70:5799-5806), and HIV specific CD8+ cytotoxic T lymphocytes (CTL) have been shown to be important in viral clearance and in the control of initial HIV-1 spread (Borrow, P et al., 1994, J. Virol., 68:6103-6110; Yang, O et al., 1996, J. Virol. 70:5799-5806). CTL responses specific to HIV also contribute to reduction in viral load during acute and asymptomatic infection (Klein, MR, et al., 1995, J. Exp. Med. 181:1365-1372; Moss, PAH et al., 1995, Proc. Natl. Acad. Sci. USA, 92:5773-5777) and may be involved in protection against the establishment of persistent HIV infections (Rowland-Jones, SL et al., 1993, Lancet, 341:860-861; Rowland-Jones, SL et.al., 1995, Nat. Med., 1:59-64). High-frequency CTL responses to HTV-1 correlated with low viral load and slow disease progression in chronically infected individuals (Musey, L et al., 1997, N. Engl. J. Med., 337:1267-1274; Ogg, GS et al., 1998, Science, 279:2103-2106.). More compelling evidence of an antiviral effect of CD8+ cells was demonstrated in controlled studies in macaques, in which CD8+ cells were depleted in vivo using a monoclonal antibody. The viral loads in these animals increased or decreased as the CD8+ cells were depleted or reappeared, respectively (Jin, X et al., 1999, J. Exp. Med., 189:991-998; Schmitz, JE, et al., 1999, Science, 283:857-860). Therefore, induction of a CTL response specific to these proteins represents a desirable response in an HIV-1 vaccine.

HIV-1 internal structural and enzymatic proteins contain conserved domains that preserve their functions and thus exhibit less antigenic diversity that may elicit more effective CTL responses (Nixon, DF et al., 1988, Nature, 336:484-487.). Efficient and durable CTL responses require endogenous antigen synthesis and processing. Current vaccine delivery techniques include immunization with live, attenuated viruses, inactivated recombinant virus infection (Letvin, NL, 1998, Science, 280:1875-1880) or plasmid DNA expression vectors. A major obstacle in the induction of CTL responses with naked DNA or recombinant virus during development of an HIV vaccine is that the expression of HIV-1 structural and enzymatic genes is tightly regulated by the virus itself. The expression of these proteins is heavily dependent upon the existence of the Rev-responsive element (RRE) of HIV-1 in recombinant vectors (Cullen, BR, 1992, Microbiol. Rev., 56:375-394; Felber, BK et al., 1989, Proc Natl Acad Sci U S A, 86:1495-1499). Poor expression is caused by the presence of AT rich inhibitory nucleotide sequences (INS) in the gag, pol and env genes, which inhibit the nuclear export and efficient expression of unspliced HIV1 mRNAs. Early studies of DNA vaccination against HIV in mice required the inclusion of Rev in their expression vectors (Lu, S et al., 1995, Virology, 209:147-154; Okuda, K et al., 1995, AIDS Res. Hum. Retroviruses, 11:933-943; Wang, B et al., 1993, Proc. Natl. Acad. Sci. U. S. A 90:4156-4160), but modification of INS has been shown to facilitate Revindependent expression of HIV-1 Gag (Qiu, J-T et al., 1999, J. Virol., 73:9145-9152; zur Megede, J et al., 2000, J Virol 74:2628-2635), allowing detectable humoral and CTL responses against this protein (Qiu, J-T et al., 1999, J. Virol., 73:9145-9152). These modified HIV-1 Gag genes produced viral-like particles of the expected density and morphology and induced an immune response to HIV-1 Gag after DNA immunization in mice (zur Megede, J et al., 2000, J Virol, 74:2628-2635).

### Summary of the Invention

30

35

25

5

10

15

20

Protective immunity against human immunodeficiency virus-1 (HIV-1) is likely to require recognition of linear and conformation epitopes from multiple HIV antigens, and whether these responses can be elicited more effectively by virion-like structures or fused CTL epitopes was previously unknown. Herein is provided a modified HIV Env with deletions in the cleavage site, fusogenic domain, and spacing of heptad repeats 1 and 2 to expose the core protein for optimal antigen

5

10

15

20

25

presentation and recognition. Additionally we provide a Gag-Pol or Gag-Pol-Nef fusion protein that is a polyprotein designed to maximize epitope presentation. The invention extends the mutation in HIV Env, Gag, Pol, and Nef to any HIV clade or strain and to related proteins of other viruses. Different combinations, different orders, and different variations on a theme are envisioned, the theme being to optimize presentation of epitopes that generate broad CTL and antibody responses.

More particularly, we have investigated the effect of specific mutations in human immunodeficiency virus type 1 (HIV-1) envelope (Env) on humoral and cellular immune responses after DNA vaccination. Mice were injected with plasmid expression vectors encoding HIV Env with modifications of conserved glycosylation sites or different COOH-terminal mutations intended to mimic a fusion intermediate. Elimination of conserved glycosylation sites did not substantially enhance humoral or CTL immunity. In contrast, a modified gp140 with deletions in the cleavage site, fusogenic domain and spacing of heptad repeats 1 and 2 enhanced humoral immunity without reducing the efficacy of the CTL response. Because of its ability to stimulate the antibody response to native gp160 without affecting cellular immunity, this modified gp140 or a related derivative is envisioned to be a useful component of an AIDS vaccine.

In addition, we have examined the immune response to HIV-1 Gag and Pol after plasmid DNA immunization with Rev-independent expression vectors encoding various forms of these proteins. Immune responses were analyzed after vaccination with four expression vectors, including Gag alone or Gag-Pol, both of which gave rise to virion-like particles (VLPs), compared to Pol alone or a Gag-Pol fusion protein that did not form VLPs. The Gag-Pol fusion protein induced the most broad and potent CTL responses to Gag and Pol in DNA-vaccinated mice, and this immunogen also readily elicited an antibody response to HIV-1 Gag and Pol determinants. Through its ability to induce broad CTL and antibody responses, this Gag-Pol fusion protein or a related derivative is envisioned to be a useful component of an AIDS vaccine.

5

1.0

15

20

25

30

## **Brief Description of the Drawings**

# Figure A. Comparison of GP2 with the structures of viral and cellular membrane fusion proteins.

(A) Recombinant Ebola Zaire GP2, (B) Recombinant Mo-55 from the TM subunit of MoMuLv, (C) Low pH-treated HA2 from influenza virus, (D) Recombinant, proteolysis-resistant core of HIV-1 gp41, (E) Recombinant SIV gp41, NMR structure, (F) Recombinant core coiled segments of the SNARES syntaxin 1-A, synaptobrevin-II, and SNAP-25B. Weissenhorn et al., 1998, Molecular Cell, 2, 605-616.

# Figure I. Schematic representation of functional domains and mutations in HIV-1 Env glycoproteins.

Full-length envelope polyprotein, gp160, with the indicated features based on the amino acid residues of HXB2 is shown (top). Functional domains include the gp120/gp41 cleavage site (residues 510/511), the fusion domain (512-527), the two heptad repeats (546-579 and 628-655), the transmembrane domain (684-705), and the cytoplasmic domain (706-856). The mutant forms of the envelope proteins are shown below the structure of gp160. COOH deletions were introduced that terminate the envelope protein at positions 752, 704, or 680 to produce gp150, gp145, or gp140, respectively. Two internal deletions that removed the cleavage site, the fusion domain, and the region between the two heptad repeats were introduced into gp160, gp150, gp145, and gp140. A further deletion in the COOHterminal region at position 592 removed the second heptad repeat, the transmembrane domain, and the interspace region to produce gp128\Delta CFI. disrupt potential glycosylation sites, asparagine (N) residues at eleven positions (88, 156, 160, 197, 230, 234, 241, 262, 276, 289, and 295) were replaced with aspartic acid (D) residues in both gp160 and gp150. Versions of both gp160 and gp150 were created with a total of 17 mutated glycosylation sites by including six additional N to D substitutions at positions 332, 339, 356, 386, 392, and 448.

## Figure II. Comparison of the expression of the HIV-1 gp160 with codonoptimized gp160.

A. Expression of plasmids encoding Rev-dependent and Rev-independent codon-modified gp160. Upper panel: expression of Rev-dependent viral gp160 (left) and codon-modified gp160 (right) in transfected 293 cells. Lower panel: comparable expression of β-actin in these transfected cells.

B-D. Expression of mutant CXCR4- and CCR5-tropic HIV Env glycoproteins with COOH-terminal truncations (panels A and B, respectively). CXCR4-tropic envelope proteins containing mutant glycosylation sites and mutant functional domains are shown (panels C and D, respectively). The indicated proteins were detected by immunoblotting as above. Cell lysates produced by transfection with vector containing no insert were used as controls (first lane in each panel).

# Figure III. Cytotoxicity of full-length gp160 is eliminated by deletion of the COOH-terminal cytoplasmic domain.

Cell rounding and detachment was not observed in control-transfected 293 cells (A), in contrast to full-length gp160 (B) and to a lesser extent in cells transfected with gp150 (C), in contrast to gp145 (D) or gp140 (E).

20

25

30

15

5

10

## Figure IV. Expression of soluble gp140\(Delta\text{CFI HIV-1}\) envelope variant.

Immunoprecipitation and Western blot analysis of supernatants from the indicated transfected cells.

## Figure V. Antibody response against HIV-1 envelope proteins in DNA immunized mice.

A. Comparison of the antibody response in mice immunized with gp140 (ΔCFI) or other Env plasmid expression vectors. Sera were collected 2 weeks after the last immunization and used to immunoprecipitate codon-altered gp160 from lysates of transfected 293 cells. The quantitation of the immunoprecipitated gp160 was done as described in Figure 5B. The average of the normalized data has been presented as a bar diagram.

B. Antibody responses in mice immunized with different mutant Env expression vectors. Antisera from immunized mice were diluted in IP buffer and 1

5

10

15

20

25

30

μl of each diluted serum was used to immunoprecipitate codon-altered HIV-1 gp160 from lysates of transfected 293 cells as described in Figure 3A. The gels were scanned and the intensity of the gp160 band was determined by densitometry using the program Image Quant and presented relative to the intensity of gp160 immunoprecipitated with positive control sera (rabbit anti-gp160), which was used to normalize data between experiments. These data are presented graphically to facilitate comparison among groups.

C. Antibody responses in mice immunized with gp140 or gp140 (\(\Delta\text{CFI}\)) were determined by immunoprecipitation and Western blotting. Animals received two booster doses (100 \(\mu\g)\) of the same plasmid, two weeks apart. Sera (1 \(\mu\l)\) collected 2 weeks after the last immunization was used to immunoprecipitate codon-optimized HIV-1 gp160 from lysates of transfected 293 cells containing 400 \(\mu\g) of total protein. Each lane corresponds to the sera from an animal immunized with either the control vector (lanes 1 and 2), CXCR4-tropic gp140 (lanes 3-6), or plasmid that expresses gp140 with the indicated mutant functional domains (lanes 7-10). A mouse monoclonal antibody to gp160 (HIV-1 V3 Monoclonal (IIIB-V3-13), NIH AIDS Reagent Program) was used as a positive control (lane 11).

# Figure VI. CTL response against HIV-1 envelope proteins in DNA immunized mice.

The CTL response to CXCR4-tropic Env and indicated deletion mutants is shown (A). The CTL responses to CXCR4-tropic envelope with glycosylation site and ΔCFI mutations are shown (B and C, respectively). Dependence of CTL activity on CD8 cells was shown by magnetic bead depletion using the indicated representative immunogens (D). Spleen cells were isolated from immunized mice two weeks after the final immunization and stimulated *in vitro* with irradiated cells expressing gp160 with addition of hIL2 (5 U/ml) at day 4. The cytolytic activity of the restimulated spleen cells was tested after 7 days against V3 peptide-pulsed BC10ME cells. Similar findings were observed with target cells that stably express full-length Env.

## Figure VII. Schematic representation of HIV-1 Gag-Pol expression constructs.

The protein sequences of Gag (amino acids 1-432) from HXB2 (GenBank accession number K03455) and Pol (amino acids 3-1003) from NL4-3 (GenBank

5

10

15

20

25

30

accession number M19921) were used to create a synthetic version of hGag-Pol using codons found in human cells. hGag-PolΔFSΔPr was made by modification of the frame shift site (FS) and inactivation of protease. For hPol, 432 amino acids were deleted from the NH<sub>2</sub>-terminal region of hGag-Pol and addition of an ATG codon. hGag was made by deletion of 925 amino acids from the COOH-terminal region of hGag-Pol. hGag-Pol, hGag-PolΔFSΔPr, hPol and hGag are expressed from the pNGVL-3 vector backbone.

# Figure VIII. HIV-1 Gag-Pol expression in transfected 293T cells and stably transfected CT26 and BC10 ME cells.

Cell lysates from 293T cells transfected with pCMV ΔR8.2 containing viral Gag-Pol (vGag-Pol) (Naldini, L et al., 1996, Science, 272:263-267), pNGVL-hGag, hPol, hGag-PolΔFSΔPr and hGag-Pol were separated by 4-15% gradient SDS-PAGE, transferred to nitrocellulose filters, and analyzed by immunoblotting with (A) human anti HIV-1-IgG, (B) monoclonal anti-p24, and (C) rabbit anti-RT. (D) Cell lysates from CT26 and BC10ME cells stably transduced with either hGag or hPol were analyzed with human anti HIV-1-IgG.

# Figure IX. Transmission electron microscopy of HIV-1 immature virus-like particles (VLP) produced by transfected 293T.

Cells were transfected with pNGVL-hGag 48 hours prior to harvesting and fixing (magnification 25,000X).

# Figure X. Gag or Pol specific CTL response mediated by CD8 positive cells in immunized mice.

Two weeks after mice were immunized with a control vector, hGag, hPol, hGag-PolΔFSΔPr, and hGag-Pol, splenic cells were harvested and sensitized with naïve mouse splenic cells pulsed with Gag or Pol peptides. One week later, effector cells were tested for cytolytic activity in a 5-h <sup>51</sup>Cr release assay using <sup>51</sup>Cr-labeled BC10ME target cells that were pulsed for 2 hours with either (A) HIV-1 Gag peptides, or (B) HIV-1 Pol peptides. (C) CD4+ or CD8+ lymphocytes were depleted from splenic cells of immunized mice with anti-mouse- CD4+ or CD8+ Dynal beads according to the manufacturer's instructions.

5

10

15

20

25

Figure XI. Gag or Pol specific CTL response mediated by CD8 positive cells in immunized mice using stable expressing cell lines as target cells.

Two weeks after immunization in mice, splenic cells were harvested and sensitized with naïve mouse splenic cells pulsed with Gag or Pol peptides. One week later, effector cells were tested for cytolytic activity in a 5-h <sup>51</sup>Cr release assay using <sup>51</sup>Cr-labeled BC10ME target cells expressing either (A) HIV-1 Gag or (B) Pol protein.

# Figure XII. HIV-1 p24 antibody ELISA assays, HIV-1 immunoblotting and immunoprecipitation Western blotting.

- (A) An HIV-1 p24 antibody ELISA assay was performed by coating 96-well plates with 50 μl of purified recombinant HIV-1<sub>IIIB</sub> p24 antigen at a concentration of 2 μg/ml in PBS buffer, pH 7.4.
- (B) HIV-1 immunoblotting of strips containing HIV-1 proteins were incubated with pooled mouse sera at a dilution of 1:25. Bands were visualized using the ECL western blotting detection reagent.
- (C) Immunoprecipitation and Western blotting of hPol gene-transfected 293T cell lysates three days after transfection with RIPA buffer. The pooled mouse serum was diluted with IP buffer. After adding 10 µg of the cell lysate containing HIV-1 Pol protein, the reactions were incubated overnight on a rotator at 4°C. The next day, 250 µl of Protein G and A Sepharose beads (10%V/V in IP buffer) were added, and the reactions were incubated on a rotator for 2 hours at 4°C. The reactions were washed 4x with IP buffer, re-suspended with 30 µl of 1X sample buffer, and then loaded onto SDS-PAGE. The reactions were transferred to an Immobilon P membrane, and then incubated with anti HIV-1-IgG. Bands were visualized using the ECL Western blotting detection reagent.

Table

|              | Figure                   | _  | 2  | 3  | 4                        | 5   | 9                    | 7  | 8  | 6                    | 10  | =                     | 12   | 13                    | 4   | 15                                | 16                                | 17                                | -81                                | 19                                 | 20                            | 21                           | 22                           | 23                            | 24                             | 25                             | 26                             | 27                             | 28                             | 59                             | 30                              |
|--------------|--------------------------|--|--|--|--------------------------|---|----------------------|--|--|----------------------|---|-----------------------|--|-----------------------|---|-----------------------------------|-----------------------------------|-----------------------------------|------------------------------------|------------------------------------|-------------------------------|------------------------------|------------------------------|-------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|---------------------------------|
|              | SEQ ID<br>NO             | -  | 2  | ~  | 4                        | 5   | 9                    | 7  | ∞  | 6                    | 02  | =                     | 12   | 13                    | 4   | 15                                | 91                                | 17                                | 18                                 | 61                                 | 20                            | 21                           | 22                           | 23                            | 24                             | 25                             | 26                             | 27                             | 28                             | 29                             | 30                              |
|              | Plasmid Map Name         | pVR1012x/s R5gp139-Nef delta MHC delta CD4/h     | pVR1012x/s R5gp157-Nef deltaMHC deltaCD4/h | pVR1012x/s X4gp139-Nef delta MHC dclta CD4/h | pVR1012x/s X4gp130-Nef/h | pVR1012x/sX4gp157-Nef delta MHC delta CD4/h | pVR1012x/s X4gp140/h | pVR1012x/s X4gp140(del F/CL del H IS)/h                              | pVR1012x/s X4gp128(del F/CL)/h                               | pVR1012x/s X4gp145/h | pVR1012x/s X4gp145(del F/CL del H IS)/h                               | pVR1012x/s R5gp140/h  | pVR1012x/sR5gp140(del F/CL del H IS)/h                               | pVR1012x/s R5gp145/h  | pVRI012x/sR5gp145(del F/CL del H IS)/h                                | pVRI012x/sR5gp140delCI(delCFI)/h  | pVR1012x/sR5gp140dclC2(delCF1)/h  | pVR1012x/sR5gp140delC3(delCF1)/h  | pVRI012x/sR5gp140delC4(delCFI)/h   | pVR1012x/sR5gp140delC5(delCFI)/h   | pVR1012x/s R5gp140(dCF1)dV1/h | pVR1012x/s gp140(dCFI)dV2/h  | pVR1012x/s gp140(dCF1)dV3/h  | pVR1012x/s R5gp140(dCF1)dV4/h | pVR1012x/s R5gp140(dCFI)dV12/h | pVR1012x/s RSgp140(dCFI)dV13/h | pVR1012x/s R5gp140(dCFI)dV14/h | pVR1012x/s R5gp140(dCFI)dv23/h | pVRI012x/s R5gp140(dCFI)dv24/h | pVR1012x/s R5gp140(dCFI)dv34/h | pVR1012x/s R5gp140(dCFI)dv123/h |
| Env Plasmids | Plasmid Name/Description | pVR1012x/s K5gp139-Nef (delta) MHC (delta) CD4/h |  | pVR1012x/s X4 gp139-Nef deltaMHC deltaCD4/h  | pVR1012x/s X4gp130-Nef/h | pVR1012x/s X4gp157-NefDMHCDCD4/h            |                      | pVR1012x/s X4gp140 ACF1/h OR pVR1012x/s X4gp140(del F/CL del H 1S)/h | pVR1012x/s X4gp128 ACF1 /h OR pVR1012x/s X4gp128(del F/CL)/h | pVR1012x/s X4gp145/h | pVR1012x/s X4gp145 ACF1 /h OR pVR1012x/s X4gp145(dei F/CL del H iS)/h | pVR1012x/s R5gp140 /h | pVR1012x/s R5gp140 ACF1 /h OR pVR1012x/sR5gp140(del F/CL del H IS)/h | pVR1012x/s R5gp145 /h | pVR1012x/s R5gp145 ACF1 /h OR pVR1012x/s R5gp145(dcl F/CL dcl H IS)/h | pVR1012x/s R5gp140delC1(delCF1)/h | pVR1012x/s R5gp140deiC2(deiCF1)/h | pVR1012x/s R5gp140delC3(delCFI)/h | pVR1012x/s R5gp140delC4 (delCF1)/h | pVR1012x/s R5gp140dclC5 (delCF1)/h | pVR1012x/s R5gp140(dCFI)/dV1  | pVR1012x/s R5gp140(dCF1)/dV2 | pVR1012x/s R5gp140(dCF1)/dV3 | pVR1012x/s R5gp140(dCFI)/dV4  | pVR1012x/s R5gp140(dCF1)/dV12  | pVR1012x/s R5gp140(dCF1)/dV13  | pVR1012x/s R5gp140(dCF1)/dV14  | pVR1012x/s R5gp140(dCF1)/dv23  | pVR1012x/s R5gp140(dCF1)/dv24  | pVR1012x/s R5gp140(dCFI)/dv34  | pvR1012x/s K5gp14U(dCFI)/dv123  |
|              | Plasmid                  | 2100   | 2200                                       | 2300   | 2302                     | 2400  | 2700                 | 2701   | 2702   | 2706                 | 2707  | 2800                  | 2801   | 2804                  | 2805  | 2810                              | 2811                              | 2812                              | 2813                               | 2814                               | 2820                          | 2821                         | 2822                         | 2823                          | 2824                           | 2825                           | 2826                           | 2827                           | 2828                           | 2829                           | 2830                            |

|              | Figure                   | 31                              | 32                              | 33                              | 34                               | 35                                     | 36                                     | 37                                     | 38                                     | 39                                      | 40                                      | 41                                      | 42                                      | 43                                      | 44                                      | 45                                       | 46                                       | 47                                       | 48                                       | 49  | 20                                 | SI                                 | 52                                 | 53                                 | 54                                 | 55                             | 56                             | 57                             | 28                             | 59                              | 09                              | 19                             | 62                              |
|--------------|--------------------------|---------------------------------|---------------------------------|---------------------------------|----------------------------------|--|--|--|--|---|---|---|---|---|---|--|--|--|--|---|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|---------------------------------|---------------------------------|--------------------------------|---------------------------------|
|              | SEQ ID<br>NO             | 31                              | 32                              | 33                              | 34                               | 35                                     | 36                                     | 37                                     | 38                                     | 39                                      | 40                                      | 41                                      | 42                                      | 43                                      | 44                                      | 45                                       | 46                                       | 47                                       | 48                                       | 49  | 20                                 | 15                                 | 52                                 | 53                                 | 54                                 | 55                             | 56                             | 57                             | 58                             | 59                              | 09                              | 61                             | 62                              |
| SI           | Plasmid Map Name         | pVR1012x/s R5gp140(dCF1)dv124/h | pVR1012x/s R5gp140(dCFI)dv134/h | pVR1012x/s R5gp140(dCF1)dv234/h | pVR1012x/s R5gp140(dCF1)dv1234/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv1/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv2/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv3/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv4/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv12/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv13/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv14/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv23/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv24/h | pAdApt CMV TbGH (+) R5gp140(dCF1)dv34/h | pAdApt CMV TbGH (+) RSgp140(dCFI)dv123/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv124/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv134/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv234/h | pAdApt CMV TbGH (+) R5gp140(dCFI)dv1234/h | pVR1012x/sR5gp145delC1(delCFI)/h   | pVR1012x/sR5gp145delC2(delCFI)/h   | pVR1012x/sR5gp145deiC3(delCF1)/h   | pVRI012x/sR5gp145delC4(dclCFI)/h   | pVR1012x/sR5gp145de1C5(de1CFI)/h   | pVR1012x/s R5 gp145(dCFI)dv1/h | pVR1012x/s R5gp145(dCF1)dv2/h  | pVR1012x/s R5gp145(dCFI)dv3/h  | pVR1012x/s R5gp145(dCFI)dv4/h  | pVR1012x/s R5gp145(dCFI)dv12/h  | pVR1012x/s R5gp145(dCF1)dv13/h  | pVR1012x/s R5gp145(dCFI)dv14/h | pVR1012x/s R5gp145(dCF1)dv23/h  |
| Env Plasmids | Plasmid Name/Description | pVR1012x/s R5gp140(dCF1)/dv124  | pVR1012x/s R5gp140(dCF1)/dv134  | pVR1012x/s R5gp140(dCF1)/dv234  | pVR1012x/s R5gp140(dCF1)/dV1234  | pAdApt: RSgp140(dCFI)/dV1              | pAdApt RSgp140(dCF1)/dV2               | pAdApt R5gp140(dCF1)/dV3               | pAdApt R5gp140(dCFI)/dV4               | pAdApt RSgp140(dCF1)/dV12               | pAdApt RSgp140(dCF1)/dV13               | pAdApt RSgp140(dCF1)/dV14               | pAdApt R5gp140(dCFI)/dV23               | pAdApt RSgp140(dCFI)/dV24               | pAdApt R5gp140(dCFI)/dv34               | pAdApt R5gp140(dCFI)/dV123               | pAdApt RSgp140(dCFI)/dV124               | pAdApt RSgp140(dCFI)/dV134               | pAdApt RSgp140(dCFI)/dV234               | pAdApt R5gp140(dCFI)/dV1234               | pVR1012x/s R5gp145deiC1 (delCF1)/h | pVR1012x/s R5gp145delC2 (delCFI)/h | pVR1012x/s R5gp145delC3 (delCF1)/h | pVR1012x/s R5gp145delC4 (delCF1)/h | pVRI012x/s R5gp145delC5 (delCFI)/h | pVR1012x/s R5gp145(dCF1)/h/dV1 | pVR1012x/s R5gp145(dCFI)/h/dV2 | pVRI012x/s R5gp145(dCFI)/h/dV3 | pVR1012x/s R5gp145(dCFI)/h/dV4 | pVR1012x/s R5gp145(dCFI)/h/dV12 | pVR1012x/s R5gp145(dCFI)/h/dV13 | pVR1012x/sR5gp145(dCFI)/h/dV14 | pVR1012x/s R5gp145(dCF1)/h/dV23 |
|              | Plasmid                  | 2831                            | 2832                            | 2833                            | 2834                             | 2835                                   | 2836                                   | 2837                                   | 2838                                   | 2839                                    | 2840                                    | 2841                                    | 2842                                    | 2843                                    | 2844                                    | 2845                                     | 2846                                     | 2847                                     | 2848                                     | 2849                                      | 2850                               | 2851                               | 2852                               | 2853                               | 2854                               | 2860                           | 2861                           | 2862                           | 2863                           | 2864                            | 2865                            | 2866                           | 2867                            |

| Į                     | Env Plasmids   |   |        |        |
|-----------------------|--|---|--------|--------|
|                       | Plasmid Name/Description   | Plasmid Map Name                                | SEQ ID | Figure |
| pVR1012x/s R5g        | pVR1012x/s R5gp145(dCF1)/h/dV24                                      | pVRI012x/s RSgp145(dCFI)dv24/h                  | 2 5    |        |
| pVKI0I2x/s R5gr       | pvR1012x/s R5gp145(dCFI)/h/dv34                                      | pVR1012x/s R5gp145(dCFI)dv34/h                  | 3 29   | 3 2    |
| pvR1012x/s K5g        | pVKI012x/s K3gp145(dCFI)/h/dV134                                     | pVR1012x/s R5gp145(dCFI)dv134/h                 | 59     | 59     |
| PVR1012X/S K3g        | p145(dCi1)/h/dV234   | pVR1012x/s R5gp145(dCFI)dv234/h                 | 99     | 99     |
| DVR 10124/5 P5        | pv NIVIZAS KJEP143(aC.F.1)aV1Z3/n                                    | pVR1012x/s R5gp145(dCFI)dv123/h                 | 67     | 19     |
| nVR 1012x/c R         | nVR 1012x/s R5m145(40-E1)h/4/1034                                    | pVR1012x/s R5gp145(dCFI)dv124/h                 | 89     | 89     |
| pVR1012x/s R5gn150 /h | 5001 50 /h   | pVR1012x/s R5gp145(dCFI)dv1234/h                | 69     | 69     |
| pVR1012x/s R5gp160/h  | Sgp160 /h  | pvK1012x/s K3gp150/h                            | 70     | 70     |
| pVR1012x/s X4gp150 /h | (4gp150/h  | pv Nicitzysnogliou/n                            | =      | 17     |
| pVR1012x/s            | pVR1012x/s X4gp150 ACF1/h OR pVR1012x/s X4gp150/de1 F/C1 de1 H 181/h | PV A1012.35 A48p130/n                           | 72     | 72     |
| pVR1012x/s            |  | DVN:012X3 A48p130(del F/CL del H 1S)/ h         | 22     | 73     |
| DVR1012x/s            | pVR1012x/s X4gp150 AB Agly /h  | WVP10125/2 V1 - 150 45 / 3 / 3 / 3 / 3          | 74     | 74     |
| DVR1012x/s            | pVR1012x/s X4gp160 /h  | pvikiotzws Adgptod AB (delta)giy/h              | 75     | 75     |
| DVR1012x/s            | pVR1012x/s X4gp160 ACF1 /h OR pVR1012x/s X4gp160rdel F/C1 del H 15/h | PVN:012X5 A4gp10U/n                             | 26     | 92     |
| OVR1012x/s            | pVR1012x/s X4gp160 Agly /h   | hvicitas Arghiol(dei r/CL dei H 15)/h           | 77     | 77.    |
| oVR1012x/s            | pVR1012x/s X4gp160 AB Agly /h OR pVR1012x/s X4gp160AB mut Agly /h    | DVR1012x/s X4m160 AR Dalvh                      | 8 5    | 78     |
| pVR1012x/s Nef/h      |  | nVR1012x/s Nofth                                | 6      | 79     |
| VR1012x/s             | pVR1012x/s NefDMHCDCD4/h   | DV NIO (ZA) S INCI/III                          | 8      | 80     |
| VR1012x/s             | pVRI012x/s NefDCD4/h   | WIND TO TO THE STATE OF THE CLASS               | - R    | 8      |
| VR1012x/s             | pVR1012x/s NefDMHC/h   | PVINIOLZAS INCI DELIZAÇÃO                       | 82     | 82     |
| VR1012 x              | 1s 89.6Pen 128/del F/CT 1/h  | pvki0i2X/s ner deita MHC/h                      | 83     | 83     |
| 35 Clade 89           | R5 Clade 89.6P pn 40 ACFI /h   | pvKi012 x/s 89.6Pgp128(del F/CL)/h              | 84     | 84     |
| VR1012 x/             | DVR1012 x/s 89 (Pom 145/del F/C) del H. 181/h.                       | pvKI012x/s 89.6Pgp140(del F/CL del H IS)/h      | 85     | 85     |
| VR1012 x/             | S 89 6Pan 160/h  | pvR1012 x/s 89.6Pgp145(del F/CL del H IS)/h     | 98     | 98     |
| 35 Clade C            | RS Clade C 20140 ACF1 A OR DVR 1012 V/c DS (clade) 12/01 12/01 12/01 | pVR1012 x/s 89.6Pgp160/h                        | 87     | 87     |
| VR1012x/s             | DVR1012x/s RSiciadoClom14Sidel File del 1116/h                       | PVR1012x/s CladeC(R5)gp140(dei F/CL dei H IS)/h | 88     | 88     |
| VR1012x/s             | DVR1012x/s R5m145 (Jade) (Rmmil) delictifu                           | pVR1012x/s CladeC(R5)gp145(del F/CL del H IS)/h | 68     | 89     |
| VR1012x/s             | RS(clade A) mid Andel E/Ci del District                              | pVK1012x/s R5gp145 CladeC (Brazil) delCFI/h     | 96     | 96     |
| VR1012x/s             | RS(clade A)midS(del E/Ct del 11 15)/11                               | pVR1012x/s R5gp140CladeA (dCF1)/h               | 16     | 91     |
| VR1012x/s             | pVR1012x/s R5(clade E)gn140/del F/CT del H 19/h                      | pVR1012x/s R5gp145CladeA (dCFI)/h               | 92     | 92     |
| VR1012x/s             | RS(clade E)m145(del F/Cr del H IS)A                                  | pvR1012X/S KSgp140 Clade E (dCF1)/h             | 93     | 93     |
|                       | יייייייייייייייייייייייייייייייייייייי                               | pv Ki 012x/s K5 gp145 Clade E (dCFI)/h          | 8      | 94     |
|                       |  |   | ١      | -      |

| (            | 1                | T                                       | Т  | Τ   | T  | Т   | 1  | 7   | T  | Τ   |  | Ţ,  | Т  | _   | T-   | 1   | Γ-   | т-  | _   | T  | Т-  | T  |   | Τ-   | Т   | Τ-   | τ-  | _  | т-  | т-   | _   | _  | _   |
|--------------|------------------|---|--|---|--|---|--|---|--|---|--|---|--|---|--|---|--|---|---|--|---|--|---|--|---|--|---|--|---|--|---|--|---|
|              | Figure           | 95                                      | 96   | 97  | 86   | 66  | 100  | 101   | 102  | 103   | 104  | 105   | 106  | 107   | 108  | 109   | 011  | Ξ   | 112   | 113  | 114   | 115  | 116   | 117  | -18   | 611  | 120   | 121  | 122   | 123  | 124   | 125  | 1361  |
|              | SEQ ID<br>NO     | 95                                      | 96   | 76  | 86   | 66  | 100  | 101   | 102  | 103   | 104  | 105   | 106  | 107   | 108  | 601   | 011  | E   | 112   | 113  | 114   | 115  | 116   | 117  | -118  | 119  | 120   | 121  | 122   | 123  | 124   | 125  | .76.1   |
|              | Plasmid Map Name | pVR1012x/s R5gp140 Clade C(SA) (dCF1)/h | pVR1012x/s R5gp145 CladeC(SA) (dCF1)/h                         | pVR1012x/s R5gp140CladeC (Brazil) dCFIdv1/h | pVR1012x/s RSgp140CladeC (Brazil) (dCFI)dv12/h | pVR1012x/s RSgp140CladeC (Brazil) (dCFI)dv123/h | pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv1234/h | pVR1012x/s R5gp140CladcC (Brazil) (dCFI)dv124/h | pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv13/h | pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv134/h | pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv14/h | pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv2/h | pVR1012x/s R5gp140CladeC (Brazil) (dCF1)dv23/h | pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv234/h | pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv24/h | pVR1012x/s R5gp140CladcC (Brazil) (dCFI)dv3/h | pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv34/h | pVR1012x/s RSgp140CladeC (Brazil) (dCFI)dv4/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv1/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv12/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv123/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv1234/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv124/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv13/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv134/h | pVR1012x/s R5gp145CladcC (Brazil) (dCFI)dv14/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv2/h | pVR1012x/s R5gp145CladcC (Brazil) (dCFI)dv23/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv234/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv24/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv3/h | pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv34/h | DVR1012x/s R5an145CladeC (Brazzi) (dCE1)dv.dh |
| Env Plasmids | _                |   | pVK1012x/s K3(clade C South Atrican)gp145(del F/CL del H IS)/h | pVKC1012(x/s)-gp140(dCF1)(Brazil C)/dV1     | pVRC1012(x/s)-gp14U(dCF1)(Brazil C)/dV12       | pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV123       | pVRC1012(x/s)-gp140(dCF1)(Brazil C)/dV1234       | pVRC:1012(x/s)-gp140(dCF1)(Brazzi C)/dV124      | pvRCi012(x/s)-gp140(dCFI)(Brazzi C)/dV13       | DVRC.1012(XS)-gp140(dC.F.1)(Brazil C)/dV134     | pv.rc.joj.2(xs)-gpj.40(ac.f.j)(Brazi) C/Jav.ja | 1001012(x/s)-gp140(aCr1)(brazil C)/dv2        | pv nc 1012(x/3)-gp1+0(ac-r/)(brazil c.)/av z.s | PVRC1012(x/S)-gp140(dCFI)(Brazil C)/dV234       | PVRC1012(XS)-gp140(aCr1)(sfazi) C//aV24        |   |  | PVRC1012(X/S)-gp1+0(dCF1)(BTZZI C)/dV4        |   |  | FVRC.1012(x/s)-gp143(dCF1)(Brazil C)/dV123      | FVRC.1012(X/S)-gp145(dCF1)(Brazil C)/dV1234      | PVBC1012(x/s)-gp143(aCri)(Brazil C)/dV124       | PVBC1012(VA)-BP145(UCF1)(BF2ZII C.)/47(13)     |   | PVBC1012(-/-)146/40Five (0)/1/3                | FVRC1012(X/S)-gp14-5(aCF1)(Brazil C)/dV2      | P(ACLIVIZ(AS)-BPI42(dCFI)(Brazil C)/dV23       | PVRC1012(XS)-gp143(dCF1)(Brazil C)/dV234        | PVPC1012(XS)-gp145(dCF1)(Brazz) C)/dV24        | DVPC1012(*/s)-gp142(aCF1)(Brazil C)/dV3       | DVPC1012(26) 20143(aCF1)(Brazil C)(aV34        | rvkciviz(x/s)-gp1+3(acri)(brazii c)/av4       |
| 1            | Plasmid          | 2308                                    | 5050   | 255   | 1650   | 2352  | 5353   | 5.524   | 2323   | 23.50   | 5258   | 5250  | 5550   | 0000  | 1000   | 2302  | 2303   | 2304  | 2303  | 2300   | 7000  | 9000   | 23.00   | 1753   | 23.77   | 27.55  | 27.00   | 5375   | 27.52   | 53.77  | 5378  | 2370   | 2717  |

|   | Figure           | 127                                  | 128                                    | 120                                     | 130                                    | 131                                   | 133                                  | 133                                   | 134                                  | 135                                 | 136                                  | 137                                   | 38                                     | 130                                 | 140                                  | 141                                 | 147                                 | 143                                  | 14                                    | 145                                    | 146                                 | 147                                 | 148  | 149                                 | 150                                  | 151                                   | 152                                | 153                                 | 154                                   | 155                                   | 156                                     |
|---|------------------|--------------------------------------|--|---|--|---------------------------------------|--------------------------------------|---------------------------------------|--------------------------------------|-------------------------------------|--------------------------------------|---------------------------------------|--|-------------------------------------|--------------------------------------|-------------------------------------|-------------------------------------|--------------------------------------|---------------------------------------|--|-------------------------------------|-------------------------------------|--|-------------------------------------|--------------------------------------|---------------------------------------|------------------------------------|-------------------------------------|---------------------------------------|---------------------------------------|---|
|   | SEQ ID           | 127                                  | 128                                    | 170                                     | 130                                    |                                       | 132                                  | 137                                   | 134                                  | 135                                 | 136                                  | 137                                   | 138                                    | 139                                 | 140                                  | 141                                 | 142                                 | 143                                  | 144                                   | 145                                    | 146                                 | 147                                 | 148  | 149                                 | 150                                  | 151                                   | 152                                | 153                                 | 154                                   | 155                                   | 156                                     |
|   | Plasmid Map Name | pVR1012x/s R5 gp140(dCF1) SA dv1/h   | pVR1012x/s R5 gp140(dCFI) SA. dv12/h   | pVR1012x/s R5 gp140(dCFI) SA dv123/h    | pVR1012x/s R5 gp140(dCFI) SA dv1234/h  | pVR1012x/s R5 gp140(dCFI) SA dv124/h  | pVR1012x/s R5 gp140(dCF1) SA dv13/h  | pVR1012x/s R5 gp140(dCFI) SA dv134/h  | pVR1012x/s R5 gp140(dCF1) SA dv14/h  | pVR1012x/s R5 gp140(dCFI) SA dv2/h  | pVR1012x/s R5 gp140(dCFI) SA dv23/h  | pVR1012x/s R5 gp140(dCFI) SA dv234/h  | pVR1012x/s R5 gp140(dCFI) SA dv24/h    | pVR1012x/s R5 gp140(dCFT) SA dv3/h  | pVR1012x/s R5 gp140(dCFI) SA dv34/h  | pVR1012x/s R5 gp140(dCFI) SA dv4/h  | pVR1012x/s R5 gp145(dCF1) SA dv1/h  | pVR1012x/s R5 gp145(dCFI) SA dv12/h  | pVR1012x/s R5 gp145(dCFI) SA dv123/h  | pVR1012x/s R5 gp145(dCFI) SA dv1234/h  | pVR1012x/s R5 gp145(dCFI) SA dv2/h  | pVR1012x/s RS gp145(dCFI) SA dv23/h | pVR1012x/s R5 gp145(dCFI) SA dv234/h             | pVR1012x/s R5 gp145(dCFI) SA dv24/h | pVR1012x/s R5 gp145(dCFI) SA dv3/h   | pVR1012x/s R5 gp145(dCFI) SA dv34/h   | pVR1012x/s R5 gp145(dCFI) SA dv4/h | pVR1012x/s R5 gp145(dCFI) SA dv13/h | pVR1012x/s R5 gp145(dCFI) SA dv134/h  | pVR1012x/s R5 gp145(dCFI) SA dv124/h  | pVR1012x/s R5 gp145(dCF1) SA dv14/h     |
|   | Plasmid Na       | pVR1012x/s R5(SA-C)gp140(dCFI) dV1/h | pVK1012x/s R5(SA-C )gp140(dCFI) dV12/h | pVR1012x/s R5(SA-C )gp140(dCFI) dV123/h | pVR1012x/s R5(SA-C)gp140(dCF1)dV1234/h | pVR1012x/s R5(SA-C)gp140(dCFI)dV124/h | pVR1012x/s R5(SA-C)gp140(dCFI)dV13/h | pVR1012x/s R5(SA-C)gp140(dCFI)dV134/h | pVR1012x/s R5(SA-C)gp140(dCFI)dV14/h | pVK1012x/s R5(SA-C)gp140(dCFI)dV2/h | pvK1012x/s R5(SA-C)gp140(dCFI)dV23/h | pVR1012x/s R5(SA-C)gp140(dCFI)dV234/h | ( pVRI012x/s R5(SA-C)gp140(dCFI)dV24/h | pvR1012x/s R3(SA-C)gp140(dCF1)dV3/h | pVR1012x/s R5(SA-C)gp140(dCFI)dV34/h | pVRI012x/s R5(SA-C)gp140(dCFI)dV4/h | pvkiuizx/s k3(SA-C)gp143(dCFI)dVI/h | pVKI012x/s K3(SA-C)gp145(dCFI)dV12/h | pVR1012x/s K3(SA-C)gp145(dCF1)dV123/h | pVKIU12x/s K3(SA-C)gp145(dCFI)dV1234/h | pvK1012x/s K3(SA-C)gp145(dCFI)dV2/h | pvkiolzws kajan-ujgpi45/dCFI)dV23/h | PV N. 01 2.25 N. 3 (3.24 - ) (3.24 ) 0 V. 23 4/n | PVNIVIZAS NJ(SA-C/BP143(UCF1)aV24/n | PVA.1012.VS K3(SA-C)gp143(aCr1)aV3/h | pviviolaxis Allon-C/Bp143(aCF1)aV34/n | PVRI012A3 AJ(3A-C)BP143[0CF1)aV4/n | PVN:012XS R3(3A-C)gp143(aCF1)aV13/h | 5V5101245 KO(SA-C)(BP140(UCT))UV134/N | PVNIOIZAS NO SAN-UBBITACIONI DA VIZAN | PYINIO1200 NO(500-1/84)143(0-1/1)0V14/N |
| j | Plasmid          | 5500                                 | 200                                    | 5502                                    | 5503                                   | 5504                                  | 5205                                 | 5206                                  | 5507                                 | 2000                                | 3                                    | 3510                                  | 3                                      | 2150                                | 5513                                 | 2514                                | CICC                                | 2010                                 | 200                                   | 2018                                   | 9166                                | 0755                                | 5532   | 2777                                | 755                                  | 25.25                                 | 2635                               | 5537                                | 8633                                  | 5570                                  | 1366                                    |

|       | Figure                                 | 157              | 158                   | 59                                 | 99                       | 19               | 162                  | 163                      | 164                              | 165  | 166  | 167  | -   | 8   | 169   | 170   | 171   | 172  | 173   | 174   |  |                | 176   |
|-------|--|------------------|-----------------------|------------------------------------|--------------------------|------------------|----------------------|--------------------------|----------------------------------|--|--|--|---|---|---|---|---|--|---|---|--|----------------|---|
|       |  | -                |                       | F                                  | -                        | -                |                      |                          |                                  | _  | -  | -  | -   | =   | <u> </u>  | _   | -   | -  | -   |   | -  | <u>-</u>       | +   |
|       | SEQ ID                                 | 157              | 158                   | 159                                | 160                      | 191              | 162                  | 163                      | 164                              | 165  | 186  | 167  |   | 108   | 169   | 170   | 171   | 172  | 173   | 174   |  | 2/2            | 176   |
|       | Plasmid Map Name                       | pVR1012x/s Gag/h | pVR1012x/s SIV Gag/h  | pVR1012x/s Gag-PoldeltaFSdeltaPr/h | pVR1012x/s SIV Gag-Pol/h | pVR1012x/s Pol/h | pVR1012x/s SIV Pol/h | pVR1012x/s Gag(fs)Pol/h  | pVR1012x/s Gag-Pol(fs)RT(-)IN(-) | pVK1012x/s-Gag(FS)-Pol-delta RT IN-IRES-R5 gp157-<br>Nef | pVR1012x/s Gag(delFS)Pol (delta PR delta RT delta IN)h | pVR1012x/s SIV Gag(deiFS)-Pol/h  | nVR 1012 Gar-Claster Dol (dolumb doll-by 3 12 min   | Which of Cart 9 10 (cellar A dellar) in               | pvRIUIZ Gag-A(dcirs) Pol (dcitaPR delta RT deltaIN)/h | pVR1012 Gag(delFS)Pol deltaPR deltaRT deltaIN/Nef/h | pVR1012 Gag(delFS)Pol deltaPR deltaRT deltaIN delta<br>Myr /h | pVR1012x/s Gag(delFS)Pol deltaPR deltaRT deltaIN delta<br>Myr/Nefh | pVR1012 Nef Gag(delFS)(del Myr) Pol (delta PR, delta RT delta IN)h    | pVR1012 Gag-C(del PS) Pol(delta PR dalta RT delta IN)     | aVB1012 Gog (4a) ESV (4a) M VII - ER 1 (7.1 ER 1.1 | RT delta INJ/h | pVR1012 Gag-A (del FS) Pol (delPR delRT delIN) Nefm |
| 13.7  | -+                                     | +                | pvkioizx/s SiV Gag /h | 1                                  | $\dagger$                | 1                | +                    | DVKIUIZXS AIV Gag-Pol /A | +                                | p v v o casta sp-roi-dena K. I nv-iKES-KS gp i 3/-jNet   | pVR1012x/s HIV gag-Pol AFS APR ART AIN /h              | pVR1012x/s SIV Gag-Pol AFS /h OR pVR1012 SIV Gag(delFS)Pol(delia PR delta RT delta IN)/h | pVRI012 Gag (delFS)Pol della PR delta RT delta IN/h | DVR 1012 Gas-A(delFS)Pol(delta PR delta RT delta INV) | nVB1012 Cau(Aa)BC\Da daka DD daka DT daka Maka Ch     |   |   | pVR1012 Gag (dciFS) Pol dcitaPR dcitaRT dcitaIN dcita MyrNcf/h     | pVR1012 Nef Gag (del fs) (del Myr) Poi (delta PR delta RT delta IN)/h | pVR1012 Gag-C(delFS)Pol(delta PR delta RT delta IN) Nef/h | Gag (dei fs) (dei Myr) Nef Pol APRARTAINA          |                | pVR1012 Gag Clade A (del fs)Pol(Δ PR Δ RT Δ IN)/h   |
| la la | III III III III III III III III III II | 2000             |                       | 9000                               | 100                      | 3                | 4101                 | 2027                     | 420                              | P P  | 4302   | 4303   | 4304  | 4305  | 4306  | 302   | 905+  | 4309   | 4310  | 4311  | 4312   |                | 4313  |

5

15

20

25

30

35

## **Detailed Description of the Invention**

The present invention provides modifications of HIV Env, Gag, Pol, and Nef that enhance immunogenicity for genetic immunization. Both HIV and SIV are genetically related members of the lentivirus genus of the Retroviridae family. Lentivirus isolates from humans are grouped into one of two types, designated HIV-1 and HIV-2. A classification scheme recognizes nine subtypes (clades) of HIV-1 (A through I) and five subtypes of HIV-2 (A through E). A compendium of HIV and SIV sequence information is found in a database prepared by Myers et al., Los Alamos, NM: Los Alamos National Laboratory.

## 10 Nucleic Acid Molecules

As indicated herein, nucleic acid molecules of the present invention may be in the form of RNA or in the form of DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include in vivo or in vitro RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) of a wild-type HIV gene; and DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode an ORF of a wild-type HIV polypeptide. Of course, the genetic code is well known in the art. Degenerate variants optimized for human codon usage are preferred.

In another aspect, the invention provides a nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described above. By "stringent hybridization conditions" is intended overnight incubation at 42 degree C in a solution comprising: 50% formamide, 5 times SSC (750 mM NaCl, 75

5

10

15

20

25

30

mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 times Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1 times SSC at about 65 degree C.

By a polynucleotide which hybridizes to a "portion" of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt), and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably about 30-70 nt of the reference polynucleotide.

By a portion of a polynucleotide of "at least 20 nt in length," for example, is intended 20 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide. Of course, a polynucleotide which hybridizes only to a complementary stretch of T (or U) resides, would not be included in a polynucleotide of the invention used to hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly T (or U) stretch or the complement thereof (e.g., practically any double-stranded DNA clone).

As indicated herein, nucleic acid molecules of the present invention which encode an HIV polypeptide may include, but are not limited to those encoding the amino acid sequence of the full-length polypeptide, by itself, the coding sequence for the full-length polypeptide and additional sequences, such as those encoding a leader or secretory sequence, such as a pre-, or pro- or prepro-protein sequence, the coding sequence of the full-length polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing, including splicing and polyadenylation signals, for example, ribosome binding and stability of mRNA; and additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities.

The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the HIV protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus

on a genome of an organism. Genes II, Lewin, B., ed., John Wiley & Sons, New

5

10

15

20

25

30

35

York (1985). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

Such variants include those produced by nucleotide substitutions, deletions or additions, which may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the HIV polypeptide or portions thereof. Also especially preferred in this regard are conservative substitutions.

Further embodiments of the invention include nucleic acid molecules comprising a polynucleotide having a nucleotide sequence at least 95% identical, and more preferably at least 96%, 97%, 98% or 99% identical to a nucleotide sequence encoding a polypeptide having the amino acid sequence of a wild-type HIV polypeptide or a nucleotide sequence complementary thereto.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a HIV polypeptide is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the HIV polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular nucleic acid molecule is at least 95%, 96%, 97%, 98% or 99% identical to the reference nucleotide sequence can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711). Bestfit uses the local homology algorithm of Smith and Waterman,

Advances in Applied Mathematics 2: 482-489 (1981), to find the best segment of homology between two sequences. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed.

The present application is directed to nucleic acid molecules at least 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequences shown herein in the Sequence Listing which encode a polypeptide having HIV polypeptide activity. By "a polypeptide having HIV activity" is intended polypeptides exhibiting HIV activity in a particular biological assay. For example, Env, Gag, and Pol protein activity can be measured for changes in immunological character by an appropriate immunological assay.

15

20

25

10

5

Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 95%, 96%, 97%, 98%, or 99% identical to a nucleic acid sequence shown herein in the Sequence Listing will encode a polypeptide "having HIV polypeptide activity." In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having HIV polypeptide activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid).

30

For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," Science 247:1306-1310 (1990), wherein the authors indicate that proteins are surprisingly tolerant of amino acid substitutions.

## Polypeptides and Fragments

The invention further provides a HIV polypeptide having the amino acid sequence encoded by an open reading frame (ORF) of a wild-type HIV gene, or a peptide or polypeptide comprising a portion thereof (e.g., gp120).

5

It will be recognized in the art that some amino acid sequences of the HIV polypeptides can be varied without significant effect of the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine activity.

10

Thus, the invention further includes variations of the HIV polypeptide which show substantial HIV polypeptide activity or which include regions of HIV protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and type substitutions. As indicated, guidance concerning which amino acid changes are likely to be phenotypically silent can be found in Bowie, J. U., et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," Science 247:1306-1310 (1990).

15

Thus, the fragment, derivative or analog of the polypeptide of the invention may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which additional amino acids are fused to the mature polypeptide, such as an IgG Fc fusion region peptide or leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

25

20

As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein (see Table A).

5

10

Table A

Conservative Amino Acid Substitutions

|                        | 1 2           |
|------------------------|---------------|
| Aromatic               | Phenylalanine |
|                        | Tryptophan    |
|                        | Tyrosine      |
| Ionizable: Acidic      | Aspartic Acid |
| ·                      | Glutamic Acid |
| Ionizable: Basic       | Arginine      |
|                        | Histidine     |
|                        | Lysine        |
| Nonionizable Polar     | Asparagine    |
|                        | Glutamine     |
|                        | Selenocystine |
|                        | Serine        |
|                        | Threonine     |
| Nonpolar (Hydrophobic) | Alanine       |
| ,                      | Glycine       |
|                        | Isoleucine    |
|                        | Leucine       |
|                        | Proline       |
|                        | Valine        |
| Sulfur Containing      | Cysteine      |
|                        | Methionine    |

Of course, the number of amino acid substitutions a skilled artisan would make depends on many factors, including those described above. Generally speaking, the number of amino acid substitutions for any given HIV polypeptide will not be more than 50, 40, 30, 20, 10, 5 or 3.

Amino acids in the HIV polypeptides of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, Science 244:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as changes in immunological character.

The polypeptides of the present invention are conveniently provided in an isolated form. By "isolated polypeptide" is intended a polypeptide removed from its native environment. Thus, a polypeptide produced and/or contained within a recombinant host cell is considered isolated for purposes of the present invention.

5

Also intended as an "isolated polypeptide" are polypeptides that have been purified, partially or substantially, from a recombinant host cell or a native source. For example, a recombinantly produced version of the HIV polypeptide can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988).

10

The polypeptides of the present invention include a polypeptide comprising a polypeptide shown herein in the Sequence Listing; as well as polypeptides which are at least 95% identical, and more preferably at least 96%, 97%, 98% or 99% identical to those described above and also include portions of such polypeptides with at least 30 amino acids and more preferably at least 50 amino acids.

1.5

20

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of an HIV polypeptide is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the HIV polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

30

35

25

As a practical matter, whether any particular polypeptide is at least 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence shown herein in the Sequence Listing can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for

-22-

5

10

15

20

25

30

35

instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed.

The polypeptides of the invention may be produced by any conventional means. Houghten, R. A. (1985) General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids. Proc. Natl. Acad. Sci. USA 82:5131-5135. This "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Pat. No. 4,631,211 to Houghten et al. (1986).

The present invention also relates to vectors which include the nucleic acid molecules of the present invention, host cells which are genetically engineered with the recombinant vectors, and the production of HIV polypeptides or fragments thereof by recombinant techniques.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will preferably include a translation initiating at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracycline or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such

5

10

15

20

25

30

35

as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986).

The HIV polypeptide can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography. hydroxylapatite chromatography chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be nonglycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

## Pharmaceutical Formulations, Dosages, and Modes of Administration

The compounds of the invention may be administered using techniques well known to those in the art. Preferably, compounds are formulated and administered by genetic immunization. Techniques for formulation and administration may be found in "Remington's Pharmaceutical Sciences", 18<sup>th</sup> ed., 1990, Mack Publishing Co., Easton, PA. Suitable routes may include parenteral delivery, such as

5

10

15

20

25

30

35

intramuscular, intradermal, subcutaneous, intramedullary injections, as well as, intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections, just to name a few. For injection, the compounds of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiological saline buffer.

In instances wherein intracellular administration of the compounds of the invention is preferred, techniques well known to those of ordinary skill in the art may be utilized. For example, such compounds may be encapsulated into liposomes, then administered as described above. Liposomes are spherical lipid bilayers with aqueous interiors. All molecules present in an aqueous solution at the time of liposome formation are incorporated into the aqueous interior. The liposomal contents are both protected from the external microenvironment and, because liposomes fuse with cell membranes, are effectively delivered into the cell cytoplasm.

Nucleotide sequences of the invention which are to be intracellularly administered may be expressed in cells of interest, using techniques well known to those of skill in the art. For example, expression vectors derived from viruses such as retroviruses, adenoviruses, adeno-associated viruses, herpes viruses, vaccinia viruses, polio viruses, or sindbis or other RNA viruses, or from plasmids may be used for delivery and expression of such nucleotide sequences into the targeted cell population. Methods for the construction of such expression vectors are well known. See, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY, and Ausubel et al., 1989, Current Protocols In Molecular Biology, Greene Publishing Associates and Wiley Interscience, NY.

The invention extends to the use of a plasmid for primary immunization (priming) of a host and the subsequent use of a recombinant virus, such as a retrovirus, adenovirus, adeno-associated virus, herpes virus, vaccinia virus, polio virus, or sindbis or other RNA virus, for boosting said host, and vice versa. For example, the host may be immunized (primed) with a plasmid by DNA immunization and receive a boost with the corresponding viral construct, and vice versa. Alternatively, the host may be immunized (primed) with a plasmid by DNA immunization and receive a boost with not the corresponding viral construct but a different viral construct, and vice versa.

5

10

15

20

25

30

35

With respect to HIV Env, Gag, and Pol, protein sequences of the invention may be used as therapeutics or prophylatics (as subunit vaccines) in the treatment of AIDS or HIV infection. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that includes the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (e.g., the concentration of the test compound which achieves a halfmaximal inhibition of viral infection relative to the amount of the event in the absence of the test compound) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography (HPLC).

The compounds of the invention may, further, serve the role of a prophylactic vaccine, wherein the host produces antibodies and/or CTL responses against HIV Env, Gag and Pol, which responses then preferably serve to neutralize HIV viruses by, for example, inhibiting further HIV infection. Administration of the compounds of the invention as a prophylactic vaccine, therefore, would comprise administering to a host a concentration of compounds effective in raising an immune response which is sufficient to elicit antibody and/or CTL reponses to HIV Env, Gag, and Pol, and/or neutralize HIV, by, for example, inhibiting HIV ability to infect cells. The exact concentration will depend upon the specific compound to be administered, but may be determined by using standard techniques for assaying the development of an immune response which are well known to those of ordinary skill in the art.

The compounds may be formulated with a suitable adjuvant in order to enhance the immunological response. Such adjuvants may include, but are not limited to mineral gels such as aluminum hydroxide; surface active substances such as lysolecithin, pluronic polyols, polyanions; other peptides; oil emulsions; and potentially useful human adjuvants such as BCG and Corynebacterium parvum.

Adjuvants suitable for co-administration in accordance with the present invention should be ones that are potentially safe, well tolerated and effective in people including QS-21, Detox-PC, MPL-SE, MoGM-CSF, TiterMax-G, CRL-1005, GERBU, TERamide, PSC97B, Adjumer, PG-026, GSK-1, GcMAF, Balethine, MPC-026, Adjuvax, CpG ODN, Betafectin, Alum, and MF59 (see Kim et al., 2000, Vaccine, 18: 597 and references therein).

Other contemplated adjuvants that may be administered include lectins, growth factors, cytokines and lymphokines such as alpha-interferon, gamma-interferon, platelet derived growth factor (PDGF), gCSF, gMCSF, TNF, epidermal growth factor (EGF), IL-1, IL-2, IL-4, IL-6, IL-8, IL-10 and IL-12.

For all such treatments described above, the exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p. 1).

It should be noted that the attending physician would know how to and when to terminate, interrupt, or adjust administration due to toxicity, or to organ dysfunctions. Conversely, the attending physician would also know to adjust treatment to higher levels if the clinical response were not adequate (precluding toxicity). The magnitude of an administered dose in the management of the viral infection of interest will vary with the severity of the condition to be treated and the route of administration. The dose and perhaps prime-boost regimen, will also vary according to the age, weight, and response of the individual patient. A program comparable to that discussed above may be used in veterinary medicine.

The pharmacologically active compounds of this invention can be processed in accordance with conventional methods of galenic pharmacy to produce medicinal agents for administration to patients, e.g., mammals including humans.

The compounds of this invention can be employed in admixture with conventional excipients, i.e., pharmaceutically acceptable organic or inorganic carrier substances suitable for parenteral, enteral (e.g., oral) or topical application which do not deleteriously react with the active compounds. Suitable

15

5

10

20

25

30

35

5

10

15

20

25

30 -

35

pharmaceutically acceptable carriers include but are not limited to water, salt solutions, alcohols, gum arabic, vegetable oils, benzyl alcohols, polyethylene glycols, gelatine, carbohydrates such as lactose, amylose or starch, magnesium stearate, talc, silicic acid, viscous paraffin, perfume oil, fatty acid monoglycerides and diglycerides, pentaerythritol fatty acid esters, hydroxy methylcellulose, polyvinyl pyrrolidone, etc. The pharmaceutical preparations can be sterilized and if desired mixed with auxiliary agents, e.g., lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers, coloring, flavoring and/or aromatic substances and the like which do not deleteriously react with the active compounds. They can also be combined where desired with other active agents, e.g., vitamins.

For parenteral application, which includes intramuscular, intradermal, subcutaneous, intranasal, intracapsular, intraspinal, intrasternal, and intravenous injection, particularly suitable are injectable, sterile solutions, preferably oily or aqueous solutions, as well as suspensions, emulsions, or implants, including suppositories. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

For enteral application, particularly suitable are tablets, dragees, liquids, drops, suppositories, or capsules. The pharmaceutical compositions may be prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated

5

10

15

20

25

30

35

edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate. A syrup, elixir, or the like can be used wherein a sweetened vehicle is employed.

Sustained or directed release compositions can be formulated, e.g., liposomes or those wherein the active compound is protected with differentially degradable coatings, e.g., by microencapsulation, multiple coatings, etc. It is also possible to freeze dry the new compounds and use the lyophilizates obtained, for example, for the preparation of products for injection.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g. gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

For topical application, there are employed as non-sprayable forms, viscous to semi-solid or solid forms comprising a carrier compatible with topical application and having a dynamic viscosity preferably greater than water. Suitable formulations include but are not limited to solutions, suspensions, emulsions, creams, ointments, powders, liniments, salves, aerosols, etc., which are, if desired, sterilized or mixed with auxiliary agents, e.g., preservatives, stabilizers, wetting agents, buffers or salts for influencing osmotic pressure, etc. For topical application, also suitable are sprayable aerosol preparations wherein the active ingredient, preferably in combination with a solid or liquid inert carrier material, is packaged in a squeeze bottle or in admixture with a pressurized volatile, normally gaseous propellant, e.g., a freon.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

## **Genetic Immunization**

Genetic immunization according to the present invention elicits an effective immune response without the use of infective agents or infective vectors. Vaccination techniques which usually do produce a CTL response do so through the use of an infective agent. A complete, broad based immune response is not generally exhibited in individuals immunized with killed, inactivated or subunit vaccines. The present invention achieves the full complement of immune responses in a safe manner without the risks and problems associated with vaccinations that use infectious agents.

10

5

According to the present invention, DNA or RNA that encodes a target protein is introduced into the cells of an individual where it is expressed, thus producing the target protein. The DNA or RNA is linked to regulatory elements necessary for expression in the cells of the individual. Regulatory elements for DNA include a promoter and a polyadenylation signal. In addition, other elements, such as a Kozak region, may also be included in the genetic construct.

15

The genetic constructs of genetic vaccines comprise a nucleotide sequence that encodes a target protein operably linked to regulatory elements needed for gene expression. Accordingly, incorporation of the DNA or RNA molecule into a living cell results in the expression of the DNA or RNA encoding the target protein and thus, production of the target protein.

20

25

30

When taken up by a cell, the genetic construct which includes the nucleotide sequence encoding the target protein operably linked to the regulatory elements may remain present in the cell as a functioning extrachromosomal molecule or it may integrate into the cell's chromosomal DNA. DNA may be introduced into cells where it remains as separate genetic material in the form of a plasmid. Alternatively, linear DNA which can integrate into the chromosome may be introduced into the cell. When introducing DNA into the cell, reagents which promote DNA integration into chromosomes may be added. DNA sequences which are useful to promote integration may also be included in the DNA molecule. Since integration into the chromosomal DNA necessarily requires manipulation of the chromosome, it is preferred to maintain the DNA construct as a replicating or non-replicating extrachromosomal molecule. This reduces the risk of damaging the cell by splicing into the chromosome without affecting the effectiveness of the vaccine. Alternatively, RNA may be administered to the cell. It is also contemplated to

5

10

15

20

25

30

provide the genetic construct as a linear minichromosome including a centromere, telomeres and an origin of replication.

The necessary elements of a genetic construct of a genetic vaccine include a nucleotide sequence that encodes a target protein and the regulatory elements necessary for expression of that sequence in the cells of the vaccinated individual. The regulatory elements are operably linked to the DNA sequence that encodes the target protein to enable expression.

The molecule that encodes a target protein is a protein-encoding molecule which is translated into protein. Such molecules include DNA or RNA which comprise a nucleotide sequence that encodes the target protein. These molecules may be cDNA, genomic DNA, synthesized DNA or a hybrid thereof or an RNA molecule such as mRNA. Accordingly, as used herein, the terms "DNA construct", "genetic construct" and "nucleotide sequence" are meant to refer to both DNA and RNA molecules.

The regulatory elements necessary for gene expression of a DNA molecule include: a promoter, an initiation codon, a stop codon, and a polyadenylation signal. In addition, enhancers are often required for gene expression. It is necessary that these elements be operable in the vaccinated individual. Moreover, it is necessary that these elements be operably linked to the nucleotide sequence that encodes the target protein such that the nucleotide sequence can be expressed in the cells of a vaccinated individual and thus the target protein can be produced.

Initiation codons and stop codons are generally considered to be part of a nucleotide sequence that encodes the target protein. However, it is necessary that these elements are functional in the vaccinated individual.

Similarly, promoters and polyadenylation signals used must be functional within the cells of the vaccinated individual.

Examples of promoters useful to practice the present invention, especially in the production of a genetic vaccine for humans, include but are not limited to promoters from Simian Virus 40 (SV40), Mouse Mammary Tumor Virus (MMTV) promoter, Human Immunodeficiency Virus (HIV) such as the HIV Long Terminal Repeat (LTR) promoter, Moloney virus, ALV, Cytomegalovirus (CMV) such as the CMV immediate early promoter, Epstein Barr Virus (EBV), Rous Sarcoma Virus (RSV) as well as promoters from human genes such as human Actin, human Myosin, human Hemoglobin, human muscle creatine and human metalothionein.

5

10

15

20

25

30

Examples of polyadenylation signals useful to practice the present invention, especially in the production of a genetic vaccine for humans, include but are not limited to SV40 polyadenylation signals and LTR polyadenylation signals. In particular, the SV40 polyadenylation signal which is in pCEP4 plasmid (Invitrogen, San Diego Calif.), referred to as the SV40 polyadenylation signal, can be used.

In addition to the regulatory elements required for DNA expression, other elements may also be included in the DNA molecule. Such additional elements include enhancers. The enhancer may be selected from the group including but not limited to: human Actin, human Myosin, human Hemoglobin, human muscle creatine and viral enhancers such as those from CMV, RSV and EBV.

Genetic constructs can be provided with mammalian origin of replication in order to maintain the construct extrachromosomally and produce multiple copies of the construct in the cell. Plasmids pCEP4 and pREP4 from Invitrogen (San Diego, Calif.) contain the Epstein Barr virus origin of replication and nuclear antigen EBNA-1 coding region which produces high copy episomal replication without integration.

An additional element may be added which serves as a target for cell destruction if it is desirable to eliminate cells receiving the genetic construct for any reason. A herpes thymidine kinase (tk) gene in an expressible form can be included in the genetic construct. When the construct is introduced into the cell, tk will be produced. The drug gangeyclovir can be administered to the individual and that drug will cause the selective killing of any cell producing tk. Thus, a system can be provided which allows for the selective destruction of vaccinated cells.

In order to be a functional genetic construct, the regulatory elements must be operably linked to the nucleotide sequence that encodes the target protein. Accordingly, it is necessary for the initiation and termination codons to be in frame with the coding sequence.

Open reading frames (ORFs) encoding the protein of interest and another or other proteins of interest may be introduced into the cell on the same vector or on different vectors. ORFs on a vector may be controlled by separate promoters or by a single promoter. In the latter arrangement, which gives rise to a polycistronic message, the ORFs will be separated by translational stop and start signals. The presence of an internal ribosome entry site (IRES) site between these ORFs permits the production of the expression product originating from the second ORF of

5

10

15

20

25

30

35

interest, or third, etc. by internal initiation of the translation of the bicistronic or polycistronic mRNA.

According to the invention, the genetic vaccine may be administered directly into the individual to be immunized or ex vivo into removed cells of the individual which are reimplanted after administration. By either route, the genetic material is introduced into cells which are present in the body of the individual. Routes of administration include, but are not limited to, intramuscular, intraperitoneal, intradermal, subcutaneous, intravenous, intraarterially, intraoccularly and oral as well as transdermally or by inhalation or suppository. Preferred routes of administration include intramuscular, intraperitoneal, intradermal and subcutaneous injection. Genetic constructs may be administered by means including, but not limited to, traditional syringes, needleless injection devices, or microprojectile bombardment gene guns. Alternatively, the genetic vaccine may be introduced by various means into cells that are removed from the individual. Such means include, for example, ex vivo transfection, electroporation, microinjection and microprojectile bombardment. After the genetic construct is taken up by the cells, they are reimplanted into the individual. It is contemplated that otherwise nonimmunogenic cells that have genetic constructs incorporated therein can be implanted into the individual even if the vaccinated cells were originally taken from another individual.

The genetic vaccines according to the present invention comprise about 1 nanogram to about 1000 micrograms of DNA. In some preferred embodiments, the vaccines contain about 10 nanograms to about 800 micrograms of DNA. In some preferred embodiments, the vaccines contain about 0.1 to about 500 micrograms of DNA. In some preferred embodiments, the vaccines contain about 1 to about 350 micrograms of DNA. In some preferred embodiments, the vaccines contain about 25 to about 250 micrograms of DNA. In some preferred embodiments, the vaccines contain about 100 micrograms DNA.

The genetic vaccines according to the present invention are formulated according to the mode of administration to be used. One having ordinary skill in the art can readily formulate a genetic vaccine that comprises a genetic construct. In cases where intramuscular injection is the chosen mode of administration, an isotonic formulation is preferably used. Generally, additives for isotonicity can include sodium chloride, dextrose, mannitol, sorbitol and lactose. In some cases, isotonic solutions such as phosphate buffered saline are preferred. Stabilizers include

gelatin and albumin. In some embodiments, a vaso-constriction agent is added to the formulation. The pharmaceutical preparations according to the present invention are provided sterile and pyrogen free.

Genetic constructs may optionally be formulated with one or more response enhancing agents such as: compounds which enhance transfection, i.e. transfecting agents; compounds which stimulate cell division, i.e. replication agents; compounds which stimulate immune cell migration to the site of administration, i.e. inflammatory agents; compounds which enhance an immune response, i.e. adjuvants or compounds having two or more of these activities.

10

15

20

5

In one embodiment, bupivacaine, a well known and commercially available pharmaceutical compound, is administered prior to, simultaneously with or subsequent to the genetic construct. Bupivacaine and the genetic construct may be formulated in the same composition. Bupivacaine is particularly useful as a cell stimulating agent in view of its many properties and activities when administered to tissue. Bupivacaine promotes and facilitates the uptake of genetic material by the cell. As such, it is a transfecting agent. Administration of genetic constructs in conjunction with bupivacaine facilitates entry of the genetic constructs into cells. Bupivacaine is believed to disrupt or otherwise render the cell membrane more permeable. Cell division and replication is stimulated by bupivacaine. Accordingly, bupivacaine acts as a replicating agent. Administration of bupivacaine also irritates and damages the tissue. As such, it acts as an inflammatory agent which elicits migration and chemotaxis of immune cells to the site of administration. In addition to the cells normally present at the site of administration, the cells of the immune system which migrate to the site in response to the inflammatory agent can come into contact with the administered genetic material and the bupivacaine. Bupivacaine, acting as a transfection agent, is available to promote uptake of genetic material by such cells of the immune system as well.

25

In addition to bupivacaine, mepivacaine, lidocaine, procains, carbocaine, methyl bupivacaine, and other similarly acting compounds may be used as response enhancing agents. Such agents acts a cell stimulating agents which promote the uptake of genetic constructs into the cell and stimulate cell replication as well as initiate an inflammatory response at the site of administration.

30

Other contemplated response enhancing agents which may function as transfecting agents and/or replicating agents and/or inflammatory agents and which may be administered include lectins, growth factors, cytokines and lymphokines

. 35

5

10

15

20

25

30

35

such as alpha-interferon, gamma-interferon, platelet derived growth factor (PDGF), gCSF, gMCSF, TNF, epidermal growth factor (EGF), IL-1, IL-2, IL-4, IL-6, IL-8, IL-10 and IL-12 as well as collagenase, fibroblast growth factor, estrogen, dexamethasone, saponins, surface active agents such as immune-stimulating complexes (ISCOMS), Freund's incomplete adjuvant, LPS analog including monophosphoryl Lipid A (MPL), muramyl peptides, quinone analogs and vesicles such as squalene and squalane, hyaluronic acid and hyaluronidase may also be used administered in conjunction with the genetic construct. In some embodiments, combinations of these agents are co-administered in conjunction with the genetic construct. In other embodiments, genes encoding these agents are included in the same or different genetic construct(s) for co-expression of the agents.

With respect to HIV Env, Gag, and Pol nucleotide sequences of the invention, particularly through genetic immunization, may be used as therapeutics or prophylatics in the treatment of AIDS or HIV infection. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that includes the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the (e.g., the concentration of the test compound which achieves a half-maximal inhibition of viral infection relative to the amount of the event in the absence of the test compound) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography (HPLC).

5

10

15

20

25

30

The compounds (for genetic immunization) of the invention may, further, serve the role of a prophylactic vaccine, wherein the host produces antibodies and/or CTL responses against HIV Env, Gag, and Pol which responses then preferably serve to neutralize HIV viruses by, for example, inhibiting further HIV infection. Administration of the compounds of the invention as a prophylactic vaccine, therefore, would comprise administering to a host a concentration of compounds effective in raising an immune response which is sufficient to elicit antibody and/or CTL reponses to HIV Env, Gag, and Pol and/or neutralize HIV, by, for example, inhibiting HIV ability to infect cells. The exact concentration will depend upon the specific compound to be administered, but may be determined by using standard techniques for assaying the development of an immune response which are well known to those of ordinary skill in the art.

#### Env

To improve the immune response to native gp160 and to expose the core protein for optimal antigen presentation and recognition, we have analyzed the immune response to modified forms of the protein. The role of conserved N-linked glycosylation sites has been studied, and analogues of fusion intermediates have been developed. Expression vectors with deletions in the cleavage site (C), the fusion peptide (F), and the interspace (I) between the two heptad repeats were termed ACFI. Plasmid DNA vaccination has been a useful technology for the development and analysis of immunogens. This method of vaccination allows appropriate post-translational modification, proper intracellular trafficking, and antigen presentation. Direct injection of naked DNA either intramuscularly or intradermally in rodents induces immune responses, and the ability to easily modify plasmid expression vectors to express different forms of HIV envelope proteins enables rapid and systematic testing of vaccine immunogens. In this disclosure, we have analyzed the immune response to modified Env candidates expressed in plasmids with modified codons to improve gene expression. Both antibody and CTL responses were analyzed after injection of plasmid DNA into muscle. A modified gp140 DNA with improved ability to elicit antibody and CTL responses to HIV Env has now been identified that is envisioned as a prototype immunogen that can elicit broadly neutralizing antibody responses to HIV.

# Exposing the Core Protein of Viral Membrane Fusion Proteins

5

10

15

20

25

30

Described herein are modified HIV envelope proteins that improve the immune response to native gp160 and expose the core protein for optimal antigen presentation and recognition. Weissenhorn et al., Molecular Cell, 2, 605-616, 1998 proposes a core protein as a model for a fusion intermediate of viral glycoproteins, where the glycoproteins are characterized by a central triple stranded coiled coil followed by a disulfide-bonded loop that reverses the chain direction and connects to an α helix packed antiparallel to the core helices, as, for example, in the case of Ebola Zaire GP2, Murine Moloney Leukemia virus (MuMoLv) 55-residue segment of the TM subunit (Mo-55), low-pH-treated influenza HA2, protease resistant core of HIV gp41, and SIV gp41 (Figure A). Thus, the strategy for improving the immune response by exposing the protease resistant core of HIV gp41 extends to other viral membrane fusion proteins that are characterized by a central triple stranded coiled coil followed by a disulfide-bonded loop that reverses the chain direction and connects to an α helix packed antiparallel to the core helices.

The present approach involves a series of internal mutations designed to replace the cleavage site (C), the fusion domain (F), and the interspace (I) between the two heptad repeats all on a backbone of COOH-terminal truncations to expose the core protein of the viral membrane fusion protein Env, based on modified gp140 as a prototype immunogen. By replacement is meant deletions, insertions, and/or substitutions of amino acid residues. In one embodiment, deletions are meant (i.e., amino acids are deleted to create the ΔCFI mutations).

In this embodiment, the  $\Delta C$  mutation is intended to eliminate proteolysis by deleting the gp120/gp41 cleavage site that links the envelope covalently to the ectodomain by 100%, 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81,%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73,%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%,11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%.

In this embodiment, the  $\Delta F$  mutation is intended to solubilize the molecule by deleting the fusion domain by 100%, 99%, 98%, 97%, 96%, 95%, 94%, 93%,

5

10

15

20

25

30

92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81,%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73,%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%,11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%.

In this embodiment, the  $\Delta I$  mutation is intended to stabilize oligomer formation by deleting the interspace between the two heptad repeats by 100%, 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81,%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73,%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%,11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%.

In this embodiment, the COOH-terminal truncation is intended to reduce toxicity by deleting the cytoplasmic domain by 100%, 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81,%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73,%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%,11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%.

In this embodiment, optionally, the COOH-terminal truncation is extended so as to solubilize the molecule by deleting the transmembrane domain by 100% 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81,%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73,%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%,11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%.

5

10

15

20

25

30

Amino acid substitutions may encompass those of a conserved or nonconserved nature. Presumably, a non-conserved substitution of a domain would act like a deletion of the domain. Conserved amino acid substitutions constitute switching one or more amino acids with amino acids of similar charge, size, and/or hydrophobicity characteristics. Non-conserved amino acid substitutions constitute switching one or more amino acids with amino acids of dissimilar charge, size, and/or hydrophobicity characteristics. The families of amino acids include the basic amino acids (lysine, arginine, histidine); the acidic amino acids (aspartic acid, glutamic acid); the non-polar amino acids (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan); the uncharged polar amino acids (glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine); and the aromatic amino acids (phenylalanine, tryptophan, and tyrosine). One or more substitutions may be introduced to achieve the  $\Delta C$  mutation intended to eliminate proteolysis by acting like a deletion of the gp120/gp41 cleavage site to link the envelope covalently to the ectodomain, the  $\Delta F$  mutation intended to solubilize the molecule by acting like a deletion of the fusion domain, the  $\Delta I$  mutation intended to stabilize oligomer formation by acting like a deletion of the interspace between the two heptad repeats, the COOH-terminal truncation intended to reduce toxicity by acting like a deletion of the cytoplasmic domain, and, optionally, the COOHterminal truncation extended so as to solubilize the molecule by acting like a deletion of the transmembrane domain.

Amino acid insertions may constitute single amino acid residues or stretches of residues. The insertions may be made at the carboxy or amino terminal end of a domain, as well as at a position internal to the domain. Such insertions will generally range from 2 to 15 amino acids in length. One or more insertions may be introduced to achieve the  $\Delta C$  mutation intended to eliminate proteolysis by acting like a deletion of the gp120/gp41 cleavage site to link the envelope covalently to the ectodomain, the  $\Delta F$  mutation intended to solubilize the molecule by acting like a deletion of the fusion domain, the  $\Delta I$  mutation intended to stabilize oligomer formation by acting like a deletion the interspace between the two heptad repeats, the COOH-terminal truncation intended to reduce toxicity by acting like a deletion of the cytoplasmic domain, and, optionally, the COOH-terminal truncation extended so as to solubilize the molecule by acting like a deletion of the transmembrane domain.

5

10

15

20

25

30

The nucleic acids of the present invention are optionally DNA, RNA, or mRNA. Most typically, the nucleic acids are provided by recombinantly making a DNA, which is expressed in a cell as RNA and/or as mRNA. Given the strategy for making the nucleic acids of the present invention, one of skill can construct a variety of clones containing functionally equivalent nucleic acids. Cloning methodologies to accomplish these ends, and sequencing methods to verify the sequence of nucleic acids are well known in the art. Examples of appropriate cloning and sequencing techniques, and instructions sufficient to direct persons of skill through many cloning exercises are found in Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, volume 152, Academic Press, Inc., San Diego, Calif. (Berger); Sambrook, et al. (1989) Molecular Cloning--A Laboratory Manual (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, N.Y., (Sambrook); and Current Protocols in Molecular Biology, F. M. Ausubel, et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1994 Supplement) (Ausubel). Product information from manufacturers of biological reagents and experimental equipment also provide information useful in known biological methods. Such manufacturers include the SIGMA chemical company (Saint Louis, Mo.), R&D systems (Minneapolis, Minn.), Pharmacia LKB Biotechnology (Piscataway, N.J.), CLONTECH Laboratories, Inc. (Palo Alto, Calif.), Chem Genes Corp., Aldrich Chemical Company (Milwaukee, Wis.), Glen Research, Inc., GIBCO BRL Life Technologies, Inc. (Gaithersberg, Md.), Fluka Chemica-Biochemika Analytika (Fluka Chemie AG, Buchs, Switzerland), Invitrogen, San Diego, Calif., and Applied Biosystems (Foster City, Calif.), as well as many other commercial sources known to one of skill.

The nucleic acid compositions of this invention, whether RNA, cDNA, mRNA, genomic DNA, or a hybrid of the various combinations, are isolated from biological sources or synthesized in vitro. The nucleic acids of the present invention are present in transformed or transfected whole cells, in transformed or transfected cell lysates, or in a partially purified or substantially pure form.

In vitro amplification techniques suitable for amplifying sequences to provide a nucleic acid or for subsequent analysis, sequencing or subcloning are known. Examples of techniques sufficient to direct persons of skill through such in vitro amplification methods, including the polymerase chain reaction (PCR) the

5

10

15

20

25

30

ligase chain reaction (LCR), Q $\beta$ -replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA) are found in Berger, Sambrook, and Ausubel, as well as Mullis, et al., (1987) U.S. Pat. No. 4,683,202; PCR Protocols A Guide to Methods and Applications (Innis, et al. eds) Academic Press Inc. San Diego, Calif. (1990) (Innis); Arnheim & Levinson (Oct. 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3:81-94; Kwoh, et al., Proc. Natl. Acad. Sci. USA, 86:1173 (1989); Guatelli, et al., Proc. Natl. Acad. Sci. USA, 87:1874 (1990); Lomell, et al., J. Clin. Chem., 35:1826 (1989); Landegren, et al., Science, 241:1077-1080 (1988); Van Brunt, Biotechnology, 8:291-294 (1990); Wu and Wallace, Gene, 4:560 (1989); Barringer, et al., Gene, 89:117 (1990), and Sooknanan and Malek, Biotechnology, 13:563-564 (1995). Improved methods of cloning in vitro amplified nucleic acids are described in Wallace, et al., U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids (up to 40 kb) are summarized in Cheng, et al., Nature, 369:684-685 (1994) and the references therein. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase: See, Ausubel, Sambrook, Innis, and Berger, all supra.

One of skill will recognize many ways of generating alterations in a given nucleic acid construct. Such well-known methods include site-directed mutagenesis, PCR amplification using degenerate oligonucleotides, exposure of cells containing the nucleic acid to mutagenic agents or radiation, chemical synthesis of a desired oligonucleotide (e.g., in conjunction with ligation and/or cloning to generate large nucleic acids) and other well-known techniques. See, Giliman and Smith, Gene 8:81-97 (1979), Roberts, et al., Nature, 328:731-734 (1987) and Sambrook, Innis, Ausubel, Berger, and Mullis (all supra).

Most modifications to nucleic acids are evaluated by routine screening techniques in suitable assays for the desired characteristic. For instance, changes in the immunological character of encoded polypeptides can be detected by an appropriate immunological assay. For instance, changes in the cellular immunological character of the polypeptide can be detected by an appropriate antibody or CTL assay. Modifications of other properties such as nucleic acid hybridization to a complementary nucleic acid, redox or thermal stability of encoded proteins, hydrophobicity, susceptibility to proteolysis, or the tendency to aggregate are all assayed according to standard techniques.

5

10

15

20

25

30

35

A wide variety of formats and labels are available and appropriate for detection of polypeptide sequences. These include analytic biochemical methods such as spectrophotometry, radiography, electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, and the like, and various immunological methods such as fluid or gel precipitin reactions, immunodiffusion (single or double), immunoelectrophoresis, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs), western blot assays, immunofluorescent assays, and the like. Several commercially available ELISA assays for the detection of retroviral components, including Env domains, are available, allowing one of skill to detect Env in biological samples.

Similarly, the detection of the nucleic acids of the present invention proceeds by well known methods such as Southern analysis, northern analysis, gel electrophoresis, PCR, radiolabeling and scintillation counting, and affinity chromatography. Many assay formats are appropriate, including those reviewed in Tijssen (1993) Laboratory Techniques in biochemistry and molecular biology-hybridization with nucleic acid probes parts I and II, Elsevier, New York and Choo (ed) (1994) Methods In Molecular Biology Volume 33--In Situ Hybridization Protocols, Humana Press Inc., New Jersey (see also, other books in the Methods in Molecular Biology series); see especially, Chapter 21 of Choo (id.) "Detection of Virus Nucleic Acids by Radioactive and Nonisotopic in Situ Hybridization". Finally, PCR is also routinely used to detect nucleic acids in biological samples (see, Innis, supra, for a general description of PCR techniques).

In one preferred embodiment, antibodies are used to detect polypeptide sequences. Methods of producing polyclonal and monoclonal antibodies are known to those of skill in the art, and many anti-HIV antibodies are available. See, e.g., Coligan (1991) Current Protocols in Immunology, Wiley/Greene, NY; and Harlow and Lane (1989) Antibodies: A Laboratory Manual, Cold Spring Harbor Press, NY; Stites, et al. (eds.) Basic and Clinical Immunology (4th ed.), Lange Medical Publications, Los Altos, Calif., and references cited therein; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.), Academic Press, New York, N.Y.; and Kohler and Milstein, Nature, 256:495-497 (1975). Other suitable techniques for antibody preparation include selection of libraries of recombinant antibodies in phage or similar vectors. See, Huse, et al., Science, 246:1275-1281 (1989); and Ward, et al., Nature, 341:544-546 (1989). Specific monoclonal and

polyclonal antibodies and antisera will usually bind with a KD of at least about 0.1 mM, more usually at least about 1  $\mu$ M, preferably at least about 0.1  $\mu$ M or better, and most typically and preferably, 0.01  $\mu$ M or better.

# Development of HIV Env Vectors

5

10

15

To develop Env glycoprotein variants that might effectively induce humoral and cellular immunity, a series of plasmid expression vectors were generated (Fig. I). HIV Env is encoded by nucleic acid sequences that contain RNA structures that limit gene expression (Andre S et al, 1998, J Virol, 72:1497-503; Qiu JT et al., 1999, J Virol, 73:9145-52; Rana TM et al., 1999, Arch Biochem Biophys, 365:175-85; Roebuck KA, Saifuddin M, 1999, Gene Expr, 8:67-84; Romano G, et al., 1999, J Cell Biochem, 75:357-68; Schneider R et al., 1997, J Virol, 71:4892-903). These vectors were therefore synthesized using codons found in human genes that allow these structures to be eliminated without affecting the amino acid sequence. Fulllength HIV Env (gp160) was highly expressed in the absence of HIV accessory proteins at levels ≥10-fold higher than Rev-dependent viral gp 160 in transfected 293 cell by Western blot analysis (Fig. IIA), and the relevant mutant proteins were detected at the expected apparent molecular weights (Fig. IIC-D). As might be anticipated, gp160 expressed from the synthetic gene was not efficiently processed in transfected 293 cells, presumably because over-expression of gp160 saturates the cellular proteases responsible for cleavage. (Binley JM et al., 2000, J. Virol, 74:627-643.) Synthetic HIV gp160 induced toxicity in transfected cells, with cell rounding and detachment evident within 48 hours (Fig. IIIA vs B). This cytotoxicity was reduced by elimination of the COOH-terminal cytoplasmic domain. Env protein that terminated at amino acid 752 (gp150) was less cytotoxic than gp160, while the shorter proteins (gp145 and gp140) produced little or no effect (Fig. IIIC, D, and E, respectively).

25

30

20

To alter Env immunogenicity, two different approaches were explored. First the effects of glycosylation on cellular and humoral immunity were evaluated by analysis of mutants in which conserved N-linked glycosylation sites were eliminated by site-directed mutagenesis. Two sets of mutations were introduced into both gp160 and gp150 (Fig. I). The first set included eleven potential sites (Δgly11), and the second set included an additional 6 sites downstream (Δgly17). Expression studies showed that the glycosylation mutants were efficiently expressed, and the

-43-

WO 02/32943 PCT/US01/25721 .

5

10.

15

20

25

30

glycosylation mutant protein was appropriately reduced in size compared to wild type gp160 or gp150, consistent with reduced N-linked glycosylation (Fig IIC).

The second approach involved a series of internal deletions designed to stabilize and expose functional domains of the protein that might be present in an extended helical structure prior to the formation of the six-member coiled-coil structure in the hairpin intermediate (Chan DC et al., 1997, Cell, 89:263-73; Weissenhorn W et al., 1997, Nature, 387:426-30). To generate this putative prehairpin structure, the cleavage site was removed to prevent the proteolytic processing of the envelope and stabilize the protein by linking it covalently to the gp41 extracellular and/or transmembrane domain. To reduce toxicity and enhance stability, the fusion peptide domain was deleted. The heptad repeats in the envelope protein are important tertiary structure domains involved in the ability of the envelope protein to form trimers (Lu M, Kim PS, 1997, J Biomol Struct Dyn, 15:465-71). The sequence between the heptad repeats was removed to stabilize the formation of trimers and eliminate formation of the hairpin intermediate. These ΔCFI deletions were introduced into full-length gp160 and COOH-terminal truncation mutants. Though cells trasfected with vectors encoding gp140\Delta CFI, gp145ΔCFI and gp160 readily expressed these proteins (Fig. IID), only gp140ΔCFI, which lacks the transmembrane domain, was readily detected in the supernatant (Fig. IV), indicating that it can give rise to soluble antigen.

# Immunogenicity of Env Mutants After DNA Vaccination

The ability of these Env proteins to elicit an immune response was determined in mice by injection with these plasmid DNA expression vectors. Antibody responses were monitored by the ability of antisera from injected mice to immunoprecipitate wild type gp160 from cell lysates by SDS-PAGE and Western blotting (Fig. V). In some cases, antibody reactivity was also confirmed by immunofluorescence. To quantitate the antibody response, immunoprecipitation followed by Western blotting with different dilutions of immunized mouse sera was tested. The intensity of the gp160 band was determined by densitometry and standardized relative to a positive control sera used to normalize data between experiments. The approximately linear dose response of gp160 intensity with serum dilution allowed quantification of the anti-gp160 antibody response in mouse sera. Data from immunized mice showed that none of the wild type Env proteins, neither the gp160, gp150, gp145, nor gp140 COOH-terminal truncations generated

consistently high antibody responses (Fig. VA). Of these proteins, gp140 was somewhat more effective than gp160 in generating antibody responses but remained low and inconsistent. Immunization with vectors designed to express glycosylation deficient envelope proteins also did not improve the humoral response. In contrast, the ΔCFI mutants in the Env truncation vectors substantially increased anti-gp160 antibody response (Fig. VA; gp145, gp140, and gp128). The longer ΔCFI Env proteins (gp160 and gp150) did not show comparable enhancement of this response. The gp140 (ΔCFI) provided more consistent and a greater increase in antibody response than gp128ΔCFI (Fig VB). In all cases, these vectors that encoded gp140ΔCFI, but not wild type gp140, induced antibodies that were reactive to native gp160 (Fig. VC).

To determine whether these modifications of Env adversely affected CTL responses, spleen cells from immunized mice were tested for their ability to lyse relevant target cells. All mice immunized with codon-altered Env vectors, including COOH-terminal deletion mutations (Fig. VIA) or glycosylation mutants (Fig. VIB), elicited strong CTL responses directed to cell lines pulsed with HIV Env peptides. These findings were also confirmed using stably transfected cells that expressed Env. Importantly, gp140 (ΔCFI), which elicited increased antibody responses relative to the comparable wild typ Env, readily induced CTL responses to native Env, as did other ΔCFI mutants (Fig. VIC). Addition of anti-CD8 antibody inhibited cytolytic activity, as did depletion using magnetic beads coupled with anti-CD8 antibody, thus confirming a cytotoxic T cell response to these immunogens after genetic immunization (Fig. VID). This response was detectable for at least six months after immunization.

# Glycosylation and ΔCFI HIV Env Mutants

5

10

15

20

25

30

To develop DNA vaccine candidates for HIV, we developed a series of synthetic genes designed to express HIV Env mutants in human cells. In the absence of HIV regulatory proteins, these codon-altered envelope protein genes expressed well in human cells. Like other DNA vaccines, immunization with these vectors elicited strong CTL responses in mice, and antibody responses were not robust in mice immunized with wild type Env expression vectors. Mutations in highly conserved N-linked glycosylation sites did not significantly alter humoral or cellular immune response to native Env. In contrast, a mutant Env with deletions in the cleavage site, fusion domain, and a region between the heptad repeats elicited a

more potent humoral immune response and retained its ability to stimulate Env-specific CTL.

Recent reports suggest that gp160 forms trimers in vivo and the domain required for trimer formation resides in the ectodomain of the gp41 (Yang X et al., 2000, J Virol, 74:5716-25). Such trimeric forms of HIV envelope protein are likely to present different epitopes to the immune system compared to monomeric gp120. In addition to the linear epitopes in the envelope, this trimeric structure is likely to expose conformational epitopes important for B cell triggering of a relevant. antibody response. In this regard, gp140 ( $\Delta$ CFI), which induced the greatest antibody response, is released in a soluble form (Fig. IVB). In contrast, wild type Env did not elicit high titer antibody responses. The toxicity of Env in mammalian cells has been seen and could limit both the amount and duration of envelope protein expression in vivo that would affect immunogenicity. The envelope is also heavily glycosylated, and removal of partial or complete gp120 glycosylation sites has resulted in higher titers of strain-specific neutralizing antibody responses to mutant SIVs in monkeys (Binley JM et al., 1998, AIDS Res Hum Retroviruses, 14:191-8; Reitter JN et al., 1998, Nat Med, 4:679-84). Though it seemed reasonable that deglycosylation would reveal epitopes otherwise masked in the native protein, we did not observe enhanced immune reactivity by DNA vaccination using different glycosylation site mutants, both in gp160 and gp150. This difference with the previous study is likely due to the fact that DNA vaccination rather than viral infection was utilized for immunization. Though glycosylation mutants are unlikely to prove helpful with this former method of immunization, we envision that modification of glycosylation sites will be effective with other vectors or adjuvants.

25

30

5

10

15

20

HIV-1 Env is proteolytically cleaved by a cellular convertase into gp120 and gp41 (Eckert DM et al., 1999, Cell, 99:103-15). The gp41 subunit is composed of cytoplasmic, transmembrane, and ectodomain segments. The role of the ectodomain of the envelope in membrane fusion, particularly its hydrophobic glycine-rich fusion peptide, is well established. Two regions with heptad coiled-coil repeats in the ectodomain of gp41 are involved in viral fusion (Eckert DM et al., 1999, Cell, 99:103-15; Weissenhorn W et al., 1996, EMBO J, 15:1507-14). Upon fusion, these two alpha helices, connected via a disulfide-stabilized loop (Gallaher WR et al., 1989, AIDS Res Hum Retroviruses, 5:431-40; Kent KA, Robinson J, 1996, AIDS, 10 Suppl A:S107-S114; Sattentau QJ et al., 1995, Virology, 206:713-7), presumably

5

10

15

20

25

30

undergo a transient conformational change to a fusion active state. These changes allow the formation of a six-member helical hairpin intermediate structure that presumably exposes the fusion peptide at the NH<sub>2</sub>-terminus of gp41, allowing fusion to the target cell membrane (Binley J, Moore JP,[published erratum appears in Nature 1997 Sep 11;389(6647):131], 1997, Nature, 387:346-8; LaCasse R et al., 1999, Science, 283:357-62). The ΔCFI mutation was intended to eliminate cleavage of gp140, remove the unstable hydrophobic region and stabilize oligomer formation. Though detailed structural data is not yet available on this protein, these mutations apparently stabilize Env in a conformation that elicits both humoral and cellular immune responses. For example, the neutralizing epitope in the ectodomain of gp41 (Muster T et al., 1994, J Virol, 68:4031-4) is present in the series of deletions and truncations of the envelope and gp140 (ACFI) is reactive with the 2F5 neutralizing monoclonal antibody that binds to this epitope. Importantly, these immunogens also induced CTL responses to Env. Though gp128 (ΔCFI) induced slightly more potent CTL activity, gp140 (ΔCFI) was better able to elicit such responses, both to peptidepulsed cells and stably transduced target cells. Thus the enhanced humoral immune response introduced by this vaccine candidate did not appear to diminish the CTL response. Taken together, these results indicate that gp140ΔCFI serves as an improved immunogen that can more effectively elicit an antibody response against the envelope by DNA vaccination while preserving its ability to induce a CTL response.

## Gag and Pol

In this disclosure, we have prepared synthetic HIV-1 B clade Gag and Pol expression vectors that are based on human (h) codon usage. These vectors encode hGag-Pol and its derivatives, hGag, hPol and an hGag-Pol fusion protein. The synthetic Gag-Pol genes show little nucleotide homology to HIV-1 but are the same in protein sequence. The modified Gag-Pol genes were subcloned into a eukaryotic plasmid expression vector for expression and DNA immunization studies. Synthetic Gag-Pol genes allowed high level Rev-independent expression of HIV-1 Gag-Pol precursor proteins in human and mouse cell lines and induced significant cellular and humoral responses in mice. The Gag-Pol fusion protein induced the broadest responses to Gag and Pol determinants and thus is envisioned as a prototype immunogen that maximizes epitope presentation.

# Eliminating The Frame Shift Site To Create Viral Polyproteins

Described herein are HIV Gag-Pol fusion proteins encoded by a continuous open reading frame so to improve the immune response to native Gag and Pol. In some viruses, translational frame shifting is exploited during protein synthesis. Specific sequences in the RNA are required for the frame shifting. The viral RNA sequences cause ribosomal slippage so that viral proteins are produced in non-equivalent ratios. For example, during translation in HIV, the ribosomes shift reading frames to synthesize Gag precursor protein and the gag-pol fusion protein in a 20:1 ratio. The strategy here is to maximize epitope presentation by transcribing an immunogen from a continuous open reading frame by eliminating the frame shift site. Thus, the strategy for improving the immune response by the use of a HIV Gag-Pol fusion protein encoded by a continuous open reading frame extends to other viral proteins that are produced in non-equivalent ratios by virtue of translational frame shifting.

15

10

5

The present invention involves HIV Gag-Pol fusion proteins encoded by a single continuous open reading frame due to mutation of the frame shift site. The frame shift site is mutated by deletions, insertions, and/or substitutions of nucleotides to create a single continuous open reading frame. In one embodiment, deletions are meant (i.e., nucleotides are deleted to create the same open reading frame).

20

The frame shift site is a mutated frame shift to create a single continuous open reading frame. For example, a set of similar retroviral gag-pol frame shift sites are optionally made for a given fusion protein, for example, by synthesizing different gag-pol frame shift regions and cloning the sequences appropriately, or by site-directed mutagenesis of a given frame shift clone. The efficacy of the frame shift sites are assessed by measuring the production of the fusion protein. The sequence that shows the highest level of expression is a "optimized" frame shift mutation for the set assessed. Alternatively, where a particular level of expression is desired, a frame shift site from a particular set of possible frame shift sites which is closest to the desired activity level is considered to be "optimized."

30

25

Although a full length Gag sequence is preferred for use in the fusion protein of the present invention, Gag is optionally deleted of subsequences without negating a polyepitope response. For example, regions of the matrix protein (p17), regions of the capsid protein (p24), regions of p2, regions of the nucleocapsid protein (p7), regions of p1, and regions of p6 can be deleted while preserving the polyepitope

35

response. Alternatively, regions of the matrix protein (p17), regions of the capsid protein (p24), regions of p2, regions of the nucleocapsid protein (p7), regions of p1, and regions of p6 can be substituted while preserving the polyepitope response. Alternatively, regions of the matrix protein (p17), regions of the capsid protein (p24), regions of p2, regions of the nucleocapsid protein (p7), regions of p1, and regions of p6 can be interrupted by insertions while preserving the polyepitope response. Optionally, regions of the matrix protein (p17), regions of the capsid protein (p24), regions of p2, regions of the nucleocapsid protein (p7), regions of p1, and regions of p6 can be mutated to inactivate these proteins.

10

15

5

Likewise, a full length Pol sequence is preferred for use in the fusion protein of the present invention, and Pol is optionally deleted of subsequences without negating a polyepitope response. For example, regions of the protease protein, regions of the reverse transcriptase protein, and regions of the integrase protein can be deleted while preserving the polyepitope response. Alternatively, regions of the protease protein, regions of the reverse transcriptase protein, and regions of the integrase protein can be substituted while preserving the polyepitope response. Alternatively, regions of the protease protein, regions of the reverse transcriptase protein, and regions of the integrase protein can be interrupted by insertions while preserving the polyepitope response. Optionally, regions of the protease protein, regions of the reverse transcriptase protein, and regions of the integrase protein can be mutated to inactivate these enzymes.

In other embodiments, the Gag-Pol fusion proteins are co-expressed with other proteins, either as fusion proteins or separately, preferably with Env sequence proteins, such as the modified HIV Env  $\Delta$ CFI proteins described herein.

25

20

In another aspect, the invention involves chimeric nucleic acid molecules. The chimeric nucleic acid molecules of the invention typically have a retroviral gag nucleic acid sequence and a retroviral pol nucleic acid sequence in the same open reading frame due to mutation of the frame shift site. The continuous open reading frame encodes a fusion protein such as those described above.

30

The chimeric nucleic acids of the present invention are optionally DNA, RNA, or mRNA. Most typically, the chimeric nucleic acids are provided by recombinantly making a DNA, which is expressed in a cell as RNA and/or as mRNA. Given the strategy for making the chimeric nucleic acids of the present invention, one of skill can construct a variety of clones containing functionally

5

10

15

20

25

30

equivalent nucleic acids. Cloning methodologies to accomplish these ends, and sequencing methods to verify the sequence of nucleic acids are well known in the art. Examples of appropriate cloning and sequencing techniques, and instructions sufficient to direct persons of skill through many cloning exercises are found in Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, volume 152, Academic Press, Inc., San Diego, Calif. (Berger); Sambrook, et al. (1989) Molecular Cloning--A Laboratory Manual (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, N.Y., (Sambrook); and Current Protocols in Molecular Biology, F. M. Ausubel, et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1994 Supplement) (Ausubel). Product information from manufacturers of biological reagents and experimental equipment also provide information useful in known biological methods. Such manufacturers include the SIGMA chemical company (Saint Louis, Mo.), R&D systems (Minneapolis, Minn.), Pharmacia LKB Biotechnology (Piscataway, N.J.), CLONTECH Laboratories, Inc. (Palo Alto, Calif.), Chem Genes Corp., Aldrich Chemical Company (Milwaukee, Wis.), Glen Research, Inc., GIBCO BRL Life Technologies, Inc. (Gaithersberg, Md.), Fluka Chemica-Biochemika Analytika (Fluka Chemie AG, Buchs, Switzerland), Invitrogen, San Diego, Calif., and Applied Biosystems (Foster City, Calif.), as well as many other commercial sources known to one of skill.

The chimeric nucleic acid compositions of this invention, whether RNA, cDNA, mRNA, genomic DNA, or a hybrid of the various combinations, are isolated from biological sources or synthesized in vitro. The chimeric nucleic acids of the present invention are present in transformed or transfected whole cells, in transformed or transfected cell lysates, or in a partially purified or substantially pure form.

In vitro amplification techniques suitable for amplifying sequences to provide a nucleic acid or for subsequent analysis, sequencing or subcloning are known. Examples of techniques sufficient to direct persons of skill through such in vitro amplification methods, including the polymerase chain reaction (PCR) the ligase chain reaction (LCR), Qβ-replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA) are found in Berger, Sambrook, and Ausubel, as well as Mullis, et al., (1987) U.S. Pat. No. 4,683,202; PCR Protocols A Guide to Methods and Applications (Innis, et al. eds) Academic Press Inc. San Diego, Calif.

5

10

15

20

25

30

35

(1990) (Innis); Amheim & Levinson (Oct. 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3:81-94; Kwoh, et al., Proc. Natl. Acad. Sci. USA, 86:1173 (1989); Guatelli, et al., Proc. Natl. Acad. Sci. USA, 87:1874 (1990); Lomell, et al., J. Clin. Chem., 35:1826 (1989); Landegren, et al., Science, 241:1077-1080 (1988); Van Brunt, Biotechnology, 8:291-294 (1990); Wu and Wallace, Gene, 4:560 (1989); Barringer, et al., Gene, 89:117 (1990), and Sooknanan and Malek, Biotechnology, 13:563-564 (1995). Improved methods of cloning in vitro amplified nucleic acids are described in Wallace, et al., U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids (up to 40 kb) are summarized in Cheng, et al., Nature, 369:684-685 (1994) and the references therein. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. See, Ausubel, Sambrook, Innis, and Berger, all supra.

One of skill will recognize many ways of generating alterations in a given nucleic acid construct. Such well-known methods include site-directed mutagenesis, PCR amplification using degenerate oligonucleotides, exposure of cells containing the nucleic acid to mutagenic agents or radiation, chemical synthesis of a desired oligonucleotide (e.g., in conjunction with ligation and/or cloning to generate large nucleic acids) and other well-known techniques. See, Giliman and Smith, Gene 8:81-97 (1979), Roberts, et al., Nature, 328:731-734 (1987) and Sambrook, Innis, Ausubel, Berger, and Mullis (all supra).

Most modifications to nucleic acids are evaluated by routine screening techniques in suitable assays for the desired characteristic. For instance, changes in the immunological character of encoded polypeptides can be detected by an appropriate immunological assay. For instance, changes in the cellular immunological character of the polypeptide can be detected by an appropriate antibody or CTL assay. Modifications of other properties such as nucleic acid hybridization to a complementary nucleic acid, redox or thermal stability of encoded proteins, hydrophobicity, susceptibility to proteolysis, or the tendency to aggregate are all assayed according to standard techniques.

A wide variety of formats and labels are available and appropriate for detection of fusion protein sequences. These include analytic biochemical methods such as spectrophotometry, radiography, electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, and the like, and various immunological

5

10

15

20

25

30

methods such as fluid or gel precipitin reactions, immunodiffusion (single or double), immunoelectrophoresis, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs), western blot assays, immunofluorescent assays, and the like. Several commercially available ELISA assays for the detection of retroviral components, including Env domains, are available, allowing one of skill to detect Env in biological samples.

Similarly, the detection of the chimeric nucleic acids of the present invention proceeds by well known methods such as Southern analysis, northern analysis, gel electrophoresis, PCR, radiolabeling and scintillation counting, and affinity chromatography. Many assay formats are appropriate, including those reviewed in Tijssen (1993) Laboratory Techniques in biochemistry and molecular biology-hybridization with nucleic acid probes parts I and II, Elsevier, New York and Choo (ed) (1994) Methods In Molecular Biology Volume 33-In Situ Hybridization Protocols, Humana Press Inc., New Jersey (see also, other books in the Methods in Molecular Biology series); see especially, Chapter 21 of Choo (id.) "Detection of Virus Nucleic Acids by Radioactive and Nonisotopic in Situ Hybridization". Finally, PCR is also routinely used to detect nucleic acids in biological samples (see, Innis, supra, for a general description of PCR techniques).

In one preferred embodiment, antibodies are used to detect polypeptide sequences. Methods of producing polyclonal and monoclonal antibodies are known to those of skill in the art, and many anti-HIV antibodies are available. See, e.g., Coligan (1991) Current Protocols in Immunology, Wiley/Greene, NY; and Harlow and Lane (1989) Antibodies: A Laboratory Manual, Cold Spring Harbor Press, NY; Stites, et al. (eds.) Basic and Clinical Immunology (4th ed.), Lange Medical Publications, Los Altos, Calif., and references cited therein; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.), Academic Press, New York, N.Y.; and Kohler and Milstein, Nature, 256:495-497 (1975). Other suitable techniques for antibody preparation include selection of libraries of recombinant antibodies in phage or similar vectors. See, Huse, et al., Science, 246:1275-1281 (1989); and Ward, et al., Nature, 341:544-546 (1989). Specific monoclonal and polyclonal antibodies and antisera will usually bind with a KD of at least about 0.1 mM, more usually at least about 1 μM, preferably at least about 0.1 μM or better, and most typically and preferably, 0.01 μM or better.

# Expression of Synthetic HIV-1 Gag and Pol Genes

5

10

15

20

25

30

Four synthetic HIV-1 Gag-and/or Pol expression vectors, hGag-Pol, hGag-PolΔFsΔPr, hPol and hGag genes were prepared (Fig. VII). To confirm expression, the synthetic or viral Gag-Pol genes were transiently transfected into 293T cells, a human kidney-derived cell line. When cell lysates were analyzed by immunoblotting with human anti-HIV-1 IgG (Fig. VIIIA), monoclonal anti-p24 (Fig. VIIIB), and rabbit anti-RT (Fig. VIIIC), Gag p55, Pol p110 and Gag-Pol p160 precursor proteins were detected in hGag, hPol, and hGag-Pol fusion plasmids transfected 293T cells, as was expected. Mature virion proteins, p24 and RTp66, were detected in the hGag-Pol gene transfected cells (Figs. VIIIA, B and C). This might be a result of the activation of protease inside cells which was itself a result of the high-level expression of Gag and Gag-Pol protein (Karacostas, V et al., 1993, Virology, 193:661-671). The expression of Gag precursor proteins from codonaltered vectors was ≥10-fold higher than viral Gag-Pol (Fig. VIII), determined by quantitative phosphorimaging. The level of accumulated Gag-Pol fusion protein was 100-fold higher in cells transfected with hGag-Pol compared to viral Gag-Pol. Virus-like particles were released from the hGag gene transfected cells (Fig. IX), detected by transmission electron microscopy. Though such particles were observed at a lower frequency with hGag-Pol, no particles were seen in cells transfected with hGag-PolΔFsΔPr or hPol vectors. Stable expression of HIV-1 Gag and Pol proteins from codon-optimized genes in mouse CT26 and BH10ME cells was also observed (Fig. VIIID).

# Induction of HIV-1 Gag and Pol CTL Responses in Mice by DNA Vaccination

To evaluate the cellular immune response to HIV-1 Gag and Pol proteins, Balb/C female mice were injected intramuscularly with the eukaryotic expression vector plasmids containing the codon-optimized genes. Two weeks after the final vaccination, splenocytes were harvested from the immunized mice and sensitized with either Gag or Pol peptide-pulsed naïve mouse splenocytes. One week later, CTL responses were analyzed using a 5-hour chromium release assay.

CTL responses specific to HIV-1 Gag and/or Pol were first analyzed using Gag or Pol peptide-pulsed BC10ME cells, or mouse fibrosarcoma cell lines derived from B/C-N cells (Collins, JL et al., 1981, J Exp. Med., 153:89-106). Immunization with hGag, hGag-PolΔFSΔPr or hGag-Pol genes induced comparably strong CTL responses specific to Gag (Fig. XA); however, after immunization with hPol, hGag-

PolΔFSΔPr or hGag-Pol genes, only the fusion protein, hGag-PolΔFSΔPr, and hPol to a lesser extent, elicited a marked CTL response to Pol (Fig. XB). To confirm that the specific killing in the CTL assays was induced by CD8<sup>+</sup> cytotoxic T lymphocytes, CD4<sup>+</sup> or CD8<sup>+</sup> cells were depleted from sensitized splenocytes by Dynal beads (Dynal, Inc., Lake Success, NY). Depletion of CD8<sup>+</sup> cells abolished the specific lysis in the hGag-Pol ΔFSΔPr gene-immunized mice, while depletion of CD4<sup>+</sup> had little effect on lysis (Fig. XC), suggesting that CD8<sup>+</sup> lymphocytes were responsible for specific cytotoxicity.

The responses were further analyzed and confirmed with the hGag or hPol gene transduced syngeneic CT26 and BC10ME cell lines. Responses to Gag in the mice immunized with the hGag, hGag-Pol\DeltaFS\DeltaPr or hGag-Pol genes were similar when peptide-pulsed cells were used as targets in the CTL assay (Fig. XIA). Mice immunized with the hPol gene generated a specific response to HIV-1 Pol on BC10ME cell lines stably expressing Pol as target cells (Fig. XIB). The same results have been observed with CT26 cell lines. These stably transfected cell lines were therefore more sensitive as target cells than peptide-pulsed cells in the Pol CTL assays.

# Antibody Response in the Immunized Mice

5

10

15

20

25

30

Sera from mice immunized with different plasmids was analyzed with a p24 ELISA. hGag immunized mice demonstrated the highest p24 antibody titers (Fig. XIIA). Unexpectedly, hGag-Pol virus-like particles elicited the lowest levels of p24 antibody. Similar results were observed by Western blotting with pooled sera (Fig. XIIB). The HIV-1 Pol specific antibodies were not detected by a commercially available Western blotting kit (Fig. XIIB), but antibodies to Pol were detected in mice immunized with hPol and hGag-Pol\DeltaFS\DeltaPr with a more sensitive method, IP/Western blotting (Fig. XIIC). Presumably, this assay is more sensitive and better able to detect native conformational epitopes. Though such antibodies were found in mice immunized with Pol and Gag-Pol fusion proteins in this assay, minimal response was detected in the mice immunized with hGag-Pol. Though both immunogens elicited similar Gag responses, the Gag-Pol fusion protein was therefore more effective in the stimulation of CTL and antibody responses to Pol.

# **HIV Gag-Pol Fusion Proteins**

In this disclosure, HIV-1 B-clade Gag and Pol genes were modified to increase Rev-independent expression of HIV-1 Gag-Pol proteins. This modification

5

10

15

20

25

30

allowed synthesis of HIV Gag and Pol, as well as fusion proteins at levels 10- to 100-fold higher than the corresponding Rev-dependent viral gene in the absence of Rev and RRE elements. These viral proteins were recognized by standard polyclonal and monoclonal antibodies (Fig. VIII), and immature VLP were produced and released from 293T cells transfected with hGag (Fig. IX).

The immune response induced by Rev-independent Gag-Pol was directed against both Gag and Pol determinants. The hGag, hGag-Pol fusion and hGag-Pol all induced strong CTL responses specific for Gag in mice immunized with plasmid DNA, but a significant Pol response was elicited only in the mice immunized with the hGag-Pol\DeltaFS\DeltaPr or Pol alone. Because immunization with hGag-Pol gene failed to induce detectable cellular or humoral responses to HIV-1 Pol protein, these findings indicate that the Gag-Pol fusion protein induces a broader range of responses and allows delivery of an immunogen with a larger number of epitopes in a single continuous open reading frame. During viral replication, viral gag-pol produces Gag precursor protein and the gag-pol fusion protein by frame shifting in a 20:1 ratio (Wilson, W et al., 1988, Cell, 55:1159-1169). The deletion of a frame shift site in hGag-PolAFSAPr results in production of only the Gag-Pol fusion protein. Expression of Gag-Pol proteins alone in human cells is not adequate to form releasable viral particles because HIV-1 viral assembly requires Gag precursor proteins (Park, J and CD Morrow, 1992, J Virol, 66:6304-6313; Smith, AJ et al., 1993, J Virol, 67:2266-2275). The ability of hGag-PolΔFSΔPr to elicit strong Gag and Pol specific CTL responses in mice may be explained by high level expression of Gag-Pol fusion protein and its retention on the inside of cells, which could create a new condition not existing during normal viral replication and could provide sufficient proteins for antigen presentation. Moreover, the mutation in viral protease prevents viral protein from intracellular activation and reduced cellular toxicity. Overexpression of this polyprotein is also likely to affect its intracellular localization/transport and is envisioned as improving antigen presentation.

The Pol gene of HIV-1 is the precursor protein for viral protease, reverse transcriptase and integrase, which are crucial to viral replication (Kohl, NE et al., 1988, Proc. Natl. Acad. Sci. USA, 85:4686-4690.). Retroviral extracellular maturation, resulting from self-activation of protease after release produces mature RT and IN which have important functions in reverse transcription and integration respectively, for HIV-1, HIV-2, and SIV replication. The catalytic cores of these

enzymes have relatively conserved domains in order to preserve their functions and thus are envisioned as inducing cross-clade CTL responses. As early as 1988, CTLs specific for HIV-1 RT were found in blood samples from HIV-1 infected individuals (Hosmalin, A et al., 1990, Proc. Natl. Acad. Sci. USA, 87:2344-2348; Walker, BD et al., 1988, Science, 240:64-66.). Relatively strong Gag-specific CTL responses have been shown in numerous non-human primate and human studies, using DNA vaccines or a live recombinant vector containing viral Gag-Pol constructs (Evans, TG, 1999, J Infect Dis, 180:290-298; Ferrari, G, 1997, Blood, 90:2406-2416; Gorse, GJ et al., 1999, Vaccine, 18:835-849; Seth, A, 1998, Proc. Natl. Acad. Sci. USA, 95:10112-10116; Seth, A et al., 2000, J Virol, 74:2502-2509), but fewer Pol-specific CTL responses have been reported. The detection of significant CTL responses specific to Pol in our disclosure may be attributed in part to establishment of stable Pol expressing cell lines, in which codon alteration and inactivation of FS and PR in the Pol gene allow high level expression of the Pol protein without cellular toxicity. Though it remains possible that the hGag-Pol, or a combination of hGag and hPol, may exert similar effects with appropriate adjuvants or with different prime-boost regimens, the Rev-independent Gag-Pol fusion protein stimulates HIV-1 Gag and Pol specific CTL responses in mice and is envisioned as proving useful in an AIDS vaccine.

20

30

15

5

10

#### PLASMID DESCRIPTIONS

# Env Plasmids

## **VRC2100**

pVR1012x/s R5gp139-Nef (delta) MHC (delta) CD4/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The envelope protein gene from pX4gp160/h (Nabel lab #1272) was ligated in frame with the mutant Nef gene from pNefDMHCDCD4/h (Nabel lab #1278) to produce

pX4gp139-NefDMHCDCD4/h contains the first 668 amino acids from the HIV envelope glycoprotein (gp139) fused to the entire mutant Nef protein. The truncated envelope polyprotein (gp139) contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain and regions important for oligomer formation. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both MHC class I and CD4 expression, point mutations were introduced into the Nef gene from pNef/h (Nabel lab #1275). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. X4gp139-NefDMHCDCD4/h is expressed from the pVR1012x/s (Nabel lab #1267) vector backbone.

# 15 **VRC2200**

5

10

20

25

30

35

pVR1012x/s R5gp157-Nef (delta) MHC (delta) CD4/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (Nabel lab number 1272) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The envelope-Nef fusion protein expressed from pR5gp157-Nef/h contains the first 820 amino acids from the HIV envelope glycoprotein (gp157) fused to the entire mutant Nef protein. The gene for gp157 was ligated in frame with the full-length mutant Nef gene from pNefDMHCDCD4/h (Nabel lab #1278) to produce pR5gp157-NefDMHCDCD4/h. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both MHC class I and CD4 expression, point mutations

were introduced into the Nef gene from pNef/h (Nabel lab #1275). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. R5gp157-NefDMHCDCD4/h is expressed from the pVR1012x/s (Nabel lab #1267) vector backbone.

5

10

15

20

25

35

#### **VRC2300**

pVR1012x/s X4 gp139-Nef deltaMHC deltaCD4/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The envelope protein gene from pX4gp160/h (Nabel lab #1272) was ligated in frame with the mutant Nef gene from pNefDMHC/h (Nabel lab #1276) to produce pX4gp139-NefDMHC/h. The envelope-Nef fusion protein expressed from pX4gp139-NefDMHC/h contains the first 668 amino acids from the HIV envelope glycoprotein (gp139) fused to the entire mutant Nef protein. The truncated envelope polyprotein (gp139) contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain and regions important for oligomer formation. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit MHC class I expression, point mutations were introduced into the Nef gene from pNef/h (Nabel lab #1275). The resulting amino acids substitutions in pNefDMHC/h are: P69A, P72A, P75A, and P78A. X4gp139-NefDMHC/h is expressed from the pVR1012x/s (Nabel lab #1267) vector backbone

## 30 **VRC2302**

pVR1012x/s X4gp130-Nef/h

For the X4gp130/h (VRC2703) portion, the protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons

optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-lenth X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 602. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Regions important for oligomer formation may be partially functional. This X4gp130/h (VRC2703) gene was fused to Nef gene. The Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create the viral Nef gene (pVR1012-Nef)(Nabel Lab #1093. The nucleotide sequence is homologous to the viral gene, and the protein encoded is the same. The expression vector backbone is pVR1012x/s (VRC2000).

15

20

25

30

35

5

10

#### **VRC2400**

pVR1012x/s X4gp157-NefDMHCDCD4/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The envelope-Nef fusion protein expressed from pX4gp157-NefDMHCDCD4/h contains the first 820 amino acids from the HIV envelope glycoprotein (gp157) fused to the entire mutant Nef protein. The truncated envelope polyprotein (gp157) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The gene for gp157 was ligated in frame with the full-length mutant Nef gene from pNefDMHCDCD4/h (VRC3600) to produce pX4gp157-NefDMHCDCD4/h. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both

MHC class I and CD4 expression, point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. X4gp160-NefDMHCDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

### **VRC2700**

5

pVR1012x/s X4gp140/h

10 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 15 N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The full-lenth X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer 20 formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC2701**

30

35

25 pVR1012x/s X4gp140(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-lenth X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking

the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC2702**

pVR1012x/s X4gp128(del F/CL)/h

10

15

20

30

35

5

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-lenth X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 592. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# 25 <u>VRC2706</u>

pVR1012x/s X4gp145/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-lenth X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 704. The truncated envelope polyprotein contains the entire

SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

20

30

35

## **VRC2707**

pVR1012x/s X4gp145(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-lenth X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 704. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

## 25 VRC2800

pVR1012x/s R5gp140/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, 1215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were

5

replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-lenth R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

20

30

35

#### **VRC2801**

pVR1012x/s R5gp140(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-lenth R5tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC2804

25 pVR1012x/s R5gp145/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession

number M68893, again using human preferred codons). The full-lenth R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC2805**

5

15

20

25

10 pVR1012x/s R5gp145(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-lenth R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2810**

pVR1012 x/s R5gp140delC1(delCFI)/h

Constant domain 1 was deleted from gp140delCFI from amino acid 33-127 and was replaced with a NheI site.

# **VRC 2811**

pVR1012 x/s R5gp140delC2 (delCFI)/h

Constant domain 2 was deleted from gp140delCFI from amino acid 199-293, and was replaced with an NheI site

# **VRC2812**

pVR 1012 x/s R5gp140delC3 (delCFI)/h

15

Constant domain 3 was deleted from gp140delCFI from amino acid 333-380, and was replaced with an NheI site.

# **VRC2813**

20 pVR1012 x/s R5gp140delC4 (delCFI)/h

Constant domain 4 was deleted from gp140delCFI from amino acide 419-458, and was replaced with an NheI site.

# 25 **VRC2814**

pVR1012 x/s R5gp140delC5 (delCFI)/h

Constant domain 5 was deleted from gp140delCFI from amino acid 472 -498 and was replaced with an Nhel site.

30

#### **VRC 2820**

pVR1012x/s R5gp140(dCFI)/dV1

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the

gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1 loop (a.a.129 to 154) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2821**

5

10

15

25

30

35

20 pVR1012x/s R5gp140(dCFI)/dV2

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2 loop (a.a.160 to 193) and transmembrane domain.

Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 2822**

5

pVR1012x/s R5gp140(dCFI)/dV3

10 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 15 N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length 20 R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V3 loop (a.a.299 to 327) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. 25 The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

# 30 **VRC 2823**

35

pVR1012x/s R5gp140(dCFI)/dV4

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The

nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V4 loop (a.a.386 to 413) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620. have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

### **VRC 2824**

pVR1012x/s R5gp140(dCFI)/dV12

20

25

30

35

5

10

15

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V2 loops (a.a.129 to 154,and a.a.160 to 193) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required

for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 2825**

5

pVR1012x/s R5gp140(dCFI)/dV13

10 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 15 N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length 20 R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V3 loops (a.a.129 to 154, and a.a.299 to 327) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino 25 acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

30

## **VRC 2826**

pVR1012x/s R5gp140(dCFI)/dV14

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the

gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V4 loops (a.a.129 to 154, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

# 20 <u>VRC 2827</u>

5

10

15

25

30

35

pVR1012x/s R5gp140(dCFI)/dV23

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion

domain, but lacking the V2, V3 loops (a.a.160 to 193,and a.a.299 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 2828**

5

15

20

25

30

10 pVR1012x/s R5gp140(dCFI)/dV24

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2, V4 loops (a.a.160 to 193, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

1.5

20

30

35

### **VRC 2829**

pVR1012x/s R5gp140(dCFI)/dV34

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V3, V4 loops (a.a.299 to 327, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

# 25 **VRC 2830**

pVR1012x/s R5gp140(dCFI)/dV123

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were

replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V2,V3 loops (a.a. 129 to 154, a.a.160 to 193,and a.a.299 to 327) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 2831**

5

10

20

25

30

15 pVR1012x/s R5gp140(dCFI)/dV124

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V2, V4 loops (a.a. 129 to 154, a.a. 160 to 193, and a.a. 386 to 413) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for

oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2832**

5 pVR1012x/s R5gp140(dCFI)/dV134

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V3, V4 loops (a.a. 129 to 154, a.a. 299 to 327, and a.a. 386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 2833**

pVR1012x/s R5gp140(dCFI)/dV234

30

35

10

15

20

25

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L,

N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2,V3,V4 loops (a.a. 160 to 193, a.a.299 to 327,and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 2834**

5

10

15

pVR1012x/s R5gp140(dCFI)/dV1234

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, 20 GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 25 N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was 30 terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V2, V3, V4 loops (a.a. 129 to 154, a.a. 160 to 193, a.a.299 to 327, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 35 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage

(F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

20

25

30

35

# **VRC 2835**

pAdApt R5gp140(dCFI)/dV1

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1 loop (a.a.129 to 154) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

#### **VRC 2836**

pAdApt R5gp140(dCFI)/dV2

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the

protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2 loop (a.a.160 to 193) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is AdApt.

## **VRC 2837**

5

10

15

pAdApt R5gp140(dCFI)/dV3

20 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 25 N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was 30 terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V3 loop (a.a.299 to 327) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. 35 The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been

deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

# 5 **VRC 2838**

10

15

20

25

pAdApt R5gp140(dCFI)/dV4

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V4 loop (a.a.386 to 413) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

# **VRC 2839**

pAdApt R5gp140(dCFI)/dV12

30

35

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L,

N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V2 loops (a.a.129 to 154,and a.a.160 to 193) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

15

20

25

30

35

10

5

# **VRC 2840**

pAdApt R5gp140(dCFI)/dV13

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L. N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V3 loops (a.a.129 to 154, and a.a.299 to 327) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2,

from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

### **VRC 2841**

10

15

20

25

5 pAdApt R5gp140(dCFI)/dV14

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V4 loops (a.a.129 to 154, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

#### **VRC 2842**

pAdApt R5gp140(dCFI)/dV23

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an

R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2, V3 loops (a.a.160 to 193,and a.a.299 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

### 15 **VRC 2843**

5

10

20

25

30

pAdApt R5gp140(dCFI)/dV24

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2, V4 loops (a.a.160 to 193, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from

amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

### **VRC 2844**

10

15

20

25

5 pAdApt R5gp140(dCFI)/dV34

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V3, V4 loops (a.a.299 to 327, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

## **VRC 2845**

pAdApt R5gp140(dCFI)/dV123

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: P53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an

R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V2,V3 loops (a.a. 129 to 154, a.a.160 to 193,and a.a.299 to 327) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt

15

20

25

30

35

10

5

#### **VRC 2846**

pAdApt R5gp140(dCFI)/dV124

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V2, V4 loops (a.a. 129 to 154, a.a. 160 to 193, and a.a. 386 to 413) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad

(H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

# 5 **VRC 2847**

10

15

20

25

35

pAdApt R5gp140(dCFI)/dV134

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V3, V4 loops (a.a. 129 to 154, a.a. 299 to 327, and a.a. 386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

#### **VRC 2848**

pAdApt R5gp140(dCFI)/dV234

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the

protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2, V3, V4 loops (a.a. 160 to 193, a.a. 299 to 327, and a.a. 386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

### **VRC 2849**

pAdApt R5gp140(dCFI)/dV1234

20

25

30

35

5

10

15

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V2,V3,V4 loops (a.a. 129 to 154, a.a. 160 to 193, a.a.299 to 327,and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad

2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

# **VRC2850**

pVR1012 x/s R5gp145delC1 (delCFI)/h

Constant domain 1 was deleted from gp145delCFI from amino acid 33-127 and was replaced with an Nhe I site.

# **VRC2851**

pVR1012 x/s R5gp145delC2 (delCFI)/h

15

5

Constant domain 2 was deleted from gp145dCFI from amino acid 199-293 and was replaced with an Nhe I site.

### **VRC2852**

20 pVR1012 x/s R5gp145delC3 (delCFI)/h

Constant domain 3 was deleted from gp145delCFI from amino acide 333-380 and was replaced with an Nhe I site.

# 25 <u>VRC2853</u>

pVR1012 x/s R5gp145delC4 (delCFI)/h

Constant domain 4 was deleted from gp145delCFI from amino acid 419-458 and was replaced with an Nhe I site.

30

35

#### **VRC2854**

pVR1012 x/s R5gp145delC5 (delCFI)/h

Constant domain 5 was deleted from gp145delCF1 from amino acid 472-498 and was replaced with an Nhe I site.

# **VRC 2860**

pVR1012x/s R5gp145(dCFI)/h/dV1

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 10 N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic 15 version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The V1 loop (a.a. 129-154) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) 20 have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

## 25 VRC 2861

30

35

pVR1012x/s R5gp145(dCFI)/h/dV2

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with

the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The V2 loop (a.a. 160-193) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

### **VRC 2862**

pVR1012x/s R5gp145(dCFI)/h/dV3

15

20

25

10

5

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The V3 loop (a.a. 299-327) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

35

30

5

10

15

20

30

35

### **VRC 2863**

pVR1012x/s R5gp145(dCFI)/h/dV4

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V4 loop (a.a. 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2864**

25 pVR1012x/s R5gp145(dCFI)/h/dV12

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession

number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The V1, V2 loops (a.a. 129-154 and 160-193) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 2865**

5

- 10

pVR1012x/s R5gp145(dCFI)/h/dV13

15 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 20 N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated 25 after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V3 loops (a.a. 129-154 and 299-327) and Fusion and Cleavage (F/CL) 30 domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

20

30

35

#### **VRC 2866**

pVR1012x/s R5gp145(dCFI)/h/dV14

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V4 loops (a.a. 129-154 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2867**

25 pVR1012x/s R5gp145(dCFI)/h/dV23

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession

number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V2, V3 loops (a.a. 160-193 and 299-327) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2868**

5

10

pVR1012x/s R5gp145(dCFI)/h/dV24

15 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 20 N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic 25 version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V2, V4 loops (a.a. 160-193 and 386-413) and Fusion and Cleavage (F/CL) 30 domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 2869**

5

10

15

20

30

35

pVR1012x/s R5gp145(dCFI)/h/dV34

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L. N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V3, V4 loops (a.a. 299-327 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2870**

25 pVR1012x/s R5gp145(dCFI)/h/dV134

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession

number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V3, V4 loops (a.a. 129-154, 299-327 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2871**

5

10

pVR1012x/s R5gp145(dCFI)/h/dV234

15 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 20 N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic 25 version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. 30 The V2, V3, V4 loops (a.a. 160-193, 299-327 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

20

30

35

# **VRC 2872**

pVR1012x/s R5gp145(dCFI)dv123/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V2, V4 loops (a.a. 129-154, 160-193 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2873**

25 pVR1012x/s R5gp145(dCFI)/h/dV124

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession

number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V2, V4 loops (a.a. 129-154, 160-193 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 2874**

5

10

pVR1012x/s R5gp145(dCFI)/h/dV1234

15 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, 20 N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic 25 version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V2, V3, V4 loops (a.a. 129-154, 160-193, 299-327 and 386-413) and 30 Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

25

30

PCT/US01/25721

#### VRC2900

pVR1012x/s R5gp150/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-lenth R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 752. The truncated envelope polyprotein (gp150) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. The expression vector backbone is pVR1012x/s (VRC2000).

# 20 <u>VRC3000</u>

pVR1012x/s R5gp160/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

# **VRC3200**

pVR1012x/s X4gp150/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The full-lenth X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 752. The truncated envelope polyprotein (gp150) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. The expression vector backbone is pVR1012x/s (VRC2000).

PCT/US01/25721

15

20

25

30

10

5

### **VRC3201**

pVR1012x/s X4gp150(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The full-lenth X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 752. The truncated envelope polyprotein (gp150) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

PCT/US01/25721

#### **VRC3202**

pVR1012x/s X4gp150 Δgly/h.

Eukaryotic vector with humanized codons expressing the HIV envelope glycoprotein gp150 from HXB2, X4 tropic mutated in the Glycosylation sites. VRC3202 pVR1012x/s X4gp150Dgly/h The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4- tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To disrupt potential glycoslylation sites in the HIV-1 envelope proteins, point mutations were introduced into the full-length X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300). The resulting amino acids substitutions in X4gp160Dgly/h are: N88D, N156D, N160D, N197E, N230D, N234D, N241D, N276D, L288V, N289D, S291T, N295D, N332D, N339D, N356D, N386D, and N448D. The full-length X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 752. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

### **VRC3203**

pVR1012x/s X4gp150 AB Δgly/h

25

30

20

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The AB designation means the amino acids from 1-307 (XbaI to EcoR1). The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To disrupt potential glycoslylation sites in the HIV-1 envelope proteins, point mutations were introduced into the full-lenth X4-tropic version of the envelope protein gene from pX4gp160/h

(VRC3300). The resulting amino acids substitutions in X4gp160Dgly/h are: N88D, N156D, N160D, N197E, N230D, N234D, N241D, N276D, L288V, N289D, S291T, and N295D. The full-lenth X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 752. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

# **VRC3300**

10 pVR1012x/s X4gp160/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

20

25

30

15

5

# **VRC3301**

pVR1012x/s X4gp160(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. Full length SU and TM-proteins are Expressed. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

PCT/US01/25721

#### **VRC3400**

pVR1012x/s X4gp160Δgly/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To disrupt potential glycoslylation sites in the HIV-1 envelope proteins, point mutations were introduced into the full-lenth X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300). The resulting amino acids substitutions in X4gp160Dgly/h are: N88D, N156D, N160D, N197E, N230D, N234D, N241D, N276D, L288V, N289D, S291T, N295D, N332D, N339D, N356D, N386D, and N448D. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

### **VRC3401**

pVR1012x/s X4gp160AB mut Δgly/h.

20

25

30

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The AB designation means the amino acids from 1-307 (XbaI to EcoR1). The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To disrupt potential glycoslylation sites in the HIV-1 envelope proteins, point mutations were introduced into the full-lenth X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300). The resulting amino acids substitutions in X4gp160Dgly/h are: N88D, N156D, N160D, N197E, N230D, N234D, N241D, N276D, L288V, N289D, S291T, and N295D. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

### **VRC3500**

pVR1012x/s Nef/h

The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. The nucleotide sequence Nef/h shows little homology to the viral gene, but the protein encoded is the same. Nef/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

10

15

20

# **VRC3600**

pVR1012x/s NefDMHCDCD4/h

The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both MHC class I and CD4 expression point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. pNefDMHCDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

#### **VRC3700**

pVR1012x/s NefDCD4/h

25

30

35

The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit CD4 expression, point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDCD/h are: D174A and D175A. pNefDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone. VRC3700 pVR1012x/s NefDCD4/h The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit CD4 expression,

point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDCD/h are: D174A and D175A. pNefDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

# 5 <u>VRC3800</u>

pVR1012x/s NefDMHC/h

The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit MHC class I expression, point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDMHC/h are: P69A, P72A, P75A, and P78A. pNefDMHC/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

15

20

25

10

### **VRC5200**

pVR1012x/s 89.6Pgp128(del F/CL)/h

The protein sequence of the envelope polyprotein (gp160) from 89.6P (Dual-tropic, GenBank accession number u89134/LOCUS:SIU89134) was used to create a synthetic version of the gene (89.6Pgp160/h) using codons optimized for expression in human cells. The nucleotide sequence 89.6Pgp160/h shows little homology to the 89.6P gene, but the protein encoded is the same. The full-lenth 89.6P, dual-tropic version of the envelope protein gene from 89.6P gp160/h (VRC3000) was terminated after the codon for amino acid 596. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. The Fusion and Cleavage (F/CL) domains, from amino acids 508-541, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

30

35

#### **VRC5201**

pVR1012x/s 89.6Pgp140(del F/CL del H IS/h

The protein sequence of the envelope polyprotein (gp160) from 89.6P (dual-tropic, GenBank accession number u89134/locus SIU89134) was used to create a synthetic

version of the gene (Dualtropic gp160/h) using codons optimized for expression in human cells. The nucleotide sequence dualtropic gp160/h shows little homology to the 89.6P gene, but the protein encoded is the same. The full-lenth 89.6P, dual-tropic version of the envelope protein gene was terminated after the codon for amino acid 683. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 508-541, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 597-625, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

### **VRC5202**

5

10

20

25

30

15 pVR1012x/s 89.6Pgp145(del F/CL del H IS/h

The protein sequence of the envelope polyprotein (gp160) from 89.6P (dual-tropic, GenBank accession number u89134/locus SIU89134) was used to create a synthetic version of the gene (Dualtropic gp160/h) using codons optimized for expression in human cells. The nucleotide sequence dualtropic gp160/h shows little homology to the 89.6P gene, but the protein encoded is the same. The full-lenth 89.6P, dual-tropic version of the envelope protein gene was terminated after the codon for amino acid 709. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 508-541, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 597-625, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

PCT/US01/25721

#### **VRC5203**

pVR1012x/s 89.6Pgp160/h

The protein sequence of the envelope polyprotein (gp160) from 89.6P (Dual-tropic, GenBank accession number U89134/LOCUS: SIU89134) was used to create a synthetic version of the gene (dual tropic gp160/h) using codons optimized for expression in human cells. The nucleotide sequence 89.6P gp160/h shows little homology to the 89.6P gene, but the protein encoded is the same. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone

10

20

25

30

5

### **VRC5300**

pVR1012x/s R5(clade C)gp140(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from 92br025 (R5-tropic, 15 GenBank accession number U52953) was used to create a synthetic version of the gene (R5gp160/h) using codons optimized for expression in human cells. The nucleotide sequence R5gp160/h shows little homology to the 92br025 gene, but the protein encoded is the same. The full-length R5-tropic version of the envelope protein was synthesized by Operon under the name: kongene. The XbaI (18nt upstream from ATG) to BgIII (1376nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the Xbal to BglII sites of VRC2701 pVR1012x/s X4gp140(del F/CL del H IS)/h backbone. Therefore, the gene is R5 (clade C) gp160/h up to the BgIII site (1376nt from ATG) and the rest of the gene after BglII site is VRC2701 pVR1012x/s X4gp140(del F/CL del H IS)/h. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000). The full-lenth X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 680.

5

10

15

20

#### **VRC5301**

pVR1012x/s R5(clade C)gp145(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from 92br025 (R5-tropic, GenBank accessionnumber U52953) was used to create a synthetic version of the gene (R5gp160/h) using codons optimized for expression in human cells. The nucleotide sequence R5gp160/h shows little homology to the 92br025 gene, but the protein encoded is the same. The full-length R5-tropic version of the envelope protein was synthesized by Operon under the name: kongene. The XbaI (18nt upstream from ATG) to BgIII (1376nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BgIII sites of VRC2701 pVR1012x/s X4gp140(del F/CL del H IS)/h backbone. Therefore, the gene is R5 (clade C) gp160/h up to the BgIII site (1376nt from ATG) and the rest of the gene after BgIII site is VRC2707 pVR1012x/s X4gp145(del F/CL del H IS)/h. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000). The full-lenth X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 704.

25

#### **VRC5303**

pVR1012x/s R5gp145CladeC(Brazil) delCFI/h

Authentic Clade C unlike 5301 which is a hybrid between C and B.

30

35

#### **VRC 5304**

pVR1012x/s R5(clade A)gp140(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from 92rw020 (R5-tropic, GenBank accessionnumber U51283) was used to create a synthetic version of the

gene (Clade-A gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the 92rw020 gene, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1837 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 486-519, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 576-604, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

### **VRC 5305**

5

10

20

25

30

pVR1012x/s R5(clade A)gp145(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from 92rw020 (R5-tropic, GenBank accessionnumber U51283) was used to create a synthetic version of the gene (Clade-A gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the 92rw020 gene, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1912 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 486-519, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 576-604, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5306**

5

10

15

25

30

35

pVR1012x/s R5(clade E)gp140(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 93th966.8 (R5-tropic, GenBank accessionnumber U08456) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 93th966.8, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1856 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 497-530, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 588-613, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

# 20 <u>VRC 5307</u>

pVR1012x/s R5(clade E)gp145(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 93th966.8 (R5-tropic, GenBank accessionnumber U08456) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 93th966.8, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1928 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 497-530, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from

amino acids 588-613, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5308**

5 pVR1012x/s R5(clade C South African)gp140(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accessionnumber AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1833 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 487-520, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 577-605, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 5309**

pVR1012x/s R5(clade C South African)gp145(del F/CL del H IS)/h

25

30

35

10

15

20

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accessionnumber AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The Xbal (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the Xbal to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be

partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 487-520, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 577-605, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 5350**

. 2

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV1

10 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCF1) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI 15 (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1 loop(a.a.133-148) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) 20 between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

25

30

35

#### **VRC 5351**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV12

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into

the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2 loops (a.a.133-191) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) I and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## 10 **VRC 5352**

5

15

20

25

35

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV123

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2,V3 loops (a.a.130-191, 330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 5353**

30 pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV1234

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little

homology to the gene 92BR025, but the protein encoded is the same. The Xbal (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2,V3,V4 loops (a.a.130-191, 330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

### **VRC 5354**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV124

15

20

25

30

10

5

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2,V4 loops (a.a. 130-191, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

25

30

35

#### **VRC 5355**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV13

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V3 loops (a.a.133-148, 330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## 20 <u>VRC 5356</u>

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV134

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V3,V4 loops (a.a.130-148, 330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad

2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5357**

5 pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV14

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V4 loops (a.a. 130-148, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC 5358**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV2

25

30

35

20

10

15

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2 loop(a.a.154-191) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS)

between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 5359**

5

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV23

10 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The Xbal 15 (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the Xbal to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2, V3 loops (a.a. 154-191, 330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, 20 and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

25

30

35

## **VRC 5360**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV234

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into

the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2,V3,V4 loops (a.a.154-191, 330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# 10 **VRC 5361**

5

15

20

25

35

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV24

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2,V4 loops (a.a. 154-191, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5362**

30 pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV3

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little

homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V3 loop(a.a.330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 5363**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV34

15

20

25

30

10

5

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V3,V4 loops (a.a.330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 5364**

5

10

15

25

30

pVRC1012(x/s)-gp140(dCFl)(Brazil C)/dV4

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V4 loop(a.a.384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## 20 VRC 5365

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV1

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1 loop(a.a.133-148) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional.

Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5366**

10

15

20

5 PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV12

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2 loops (a.a.133-191) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5367**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV123

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1, V2,V3 loops (a.a. 133-191,330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic

domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## 5 VRC 5368

10

15

20

30

35

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV1234

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2, V3,V4 loops (a.a. 133-191,330-358,384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5369**

25 PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV124

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The Xbal (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the Xbal to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1, V2,V4 loops (a.a. 133-148,154-191,384-

408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 5370**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV13

10

15

20

5

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V3 loops (a.a.133-148, 330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

25

30

35

## **VRC 5371**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV134

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The Xbal (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into

the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1, V3,V4 loops (a.a. 133-148,330-358,384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## 10 **VRC 5372**

5

15

20

25

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV14

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V4 loops (a.a.133-148, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5373**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV2

30

35

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The Xbal

(18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2 loop (a.a.154-191) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

10

15

20

25

35

5

## **VRC 5374**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV23

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2,V3 loops (a.a.154-191,330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 5375**

30 PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV234

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little

homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2, V3,V4 loops (a.a. 154-191,330-358,384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5376**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV24

15

20

25

10

5

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2,V4 loops (a.a.154-191, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

30

35

#### **VRC 5377**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV3

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version

of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The Xbal (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the Xbal to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V3 loop (a.a.330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC 5378**

5

10

20

25

30

15 PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV34

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V3,V4 loops (a.a.330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

### **VRC 5379**

5

10

15

25

30

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV4

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V4 loop (a.a.384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains(a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC5500**

20 pVR1012x/s R5(SA-C)gp140(dCFI) dV1/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1 loop (a.a.136-150), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

25

30

35

### **VRC5501**

pVR1012x/s R5(SA-C)gp140(dCFI) dV12/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2 loops (a.a.136-194), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 5502**

20 pVR1012x/s R5(SA-C )gp140(dCFI) dV123/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2, V3 loops (a.a. 136-194, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC5503**

pVR1012x/s R5(SA-C)gp140(dCFI)dV1234/h

5 The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream 10 from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2, V3, V4 loops (a.a.136-194,297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions 15 important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## 20 VRC5504

25

30

pVR1012x/s R5(SA-C)gp140(dCFI)dV124/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2, V4 loops (a.a. 136-194, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their

Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC5505**

10

15

20

5 pVR1012x/s R5(SA-C)gp140(dCFI)dV13/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V3 loops (a.a.136-150, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC5506**

pVR1012x/s R5(SA-C)gp140(dCFI)dV134/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V3, V4 loops (a.a. 136-150, 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a. 577-605), the transmembrane domain and the intracellular region. Regions

important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# 5 <u>VRC5507</u>

pVR1012x/s R5(SA-C)gp140(dCFI)dV14/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V4 loops (a.a.136-150, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC5508**

pVR1012x/s R5(SA-C)gp140(dCFI)dV2/h

25

30

35

10

15

20

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2 loop (a.a.156-194), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane

domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## 5 **VRC5509**

pVR1012x/s R5(SA-C)gp140(dCFI)dV23/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V3 loops (a.a. 156-194, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a. 577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC5510**

pVR1012x/s R5(SA-C)gp140(dCFI)dV234/h

25

30

35

10

15

20

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V3, V4 loops (a.a. 156-194, 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2

(a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

20

30

35

### **VRC5511**

pVR1012x/s R5(SA-C)gp140(dCFI)dV24/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V4 loops (a.a.156-194, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC5512**

25 pVR1012x/s R5(SA-C)gp140(dCFI)dV3/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The Xbal (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the Xbal to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V3 loop (a.a.297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520),

the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

20

30

35

## **VRC5513**

pVR1012x/s R5(SA-C)gp140(dCFI)dV34/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V3, V4 loops (a.a. 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a. 577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC5514**

25 pVR1012x/s R5(SA-C)gp140(dCFI)dV4/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V4 loop (a.a 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520),

the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

10

15

20

### **VRC5515**

pVR1012x/s R5(SA-C)gp145(dCFl)dV1/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1 loop (a.a.136-150), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC5516**

pVR1012x/s R5(SA-C)gp145(dCFI)dV12/h

25

30

35

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The Xbal (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the Xbal to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1,V2 loops (a.a.136-194), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the

intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# 5 <u>VRC5517</u>

10

15

.20

pVR1012x/s R5(SA-C)gp145(dCFI)dV123/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2, V3 loops (a.a.136-194, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC5518**

pVR1012x/s R5(SA-C)gp145(dCFI)dV1234/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1,V2,V3,V4 loops (a.a.136-194, 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer

formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# 5 <u>VRC5519</u>

10

1.5

20

pVR1012x/s R5(SA-C)gp145(dCFI)dV2/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2 loop (a.a.156-194), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC5520**

pVR1012x/s R5(SA-C)gp145(dCFI)dV23/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V3 loops (a.a.156-194, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially

functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## VRC5521

10

15

20

5 pVR1012x/s R5(SA-C)gp145(dCFI)dV234/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V3, V4 loops (a.a.156-194, 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC5522**

pVR1012x/s R5(SA-C)gp145(dCFI)dV24/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V4 loops (a.a.156-194, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially

functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC5523**

5 pVR1012x/s R5(SA-C)gp145(dCFI)dV3/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V3 loop (a.a.297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

20

25

30

10

15

#### **VRC5524**

pVR1012x/s R5(SA-C)gp145(dCFI)dV34/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V3, V4 loops (a.a. 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially

functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC5525**

5 pVR1012x/s R5(SA-C)gp145(dCFI)dV4/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V4 loop (a.a.385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

20

25

30

10

15

#### **VRC5526**

pVR1012x/s R5(SA-C)gp145(dCFI)dV13/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V3 loops (a.a.136-150, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially

functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# **VRC5527**

10

15

20

5 pVR1012x/s R5(SA-C)gp145(dCFI)dV134/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V3, V4 loops (a.a.136-150, 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

## **VRC5528**

pVR1012x/s R5(SA-C)gp145(dCFI)dV124/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2, V4 loops (a.a.136-150, 156-194, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer

formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# 5 **VRC5529**

pVR1012x/s R5(SA-C)gp145(dCFI)dV14/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V4 loops (a.a.136-150, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

# Gag-Pol Plasmids

#### **VRC3900**

pVR1012x/s Gag/h

25

30

10

15

20

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein to produce pGag(fs)Pol/h (VRC4200). The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene using codons optimized for expression in human cells. To produce a gene that

expresses all of the gag proteins, the region encoding pol amino acids 77-1003 were deleted from Gag(fs)Pol/h to produce Gag/h. This construct also encodes most of the protease (98 amino acids) gene encoding amino acids 3-77. Gag/h is expressed from the pVR1012x/s vector backbone.

5

10

# **VRC3901**

pVR1012x/s SIV Gag/h

The protein sequence of the gag polyprotein (amino acid from 1-550) SIVmac239(GenBank accession number M33262) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the SIVMac239 gene, but the protein encoded is the same.

## 15 **VRC4000**

pVR1012x/s Gag-Pol/h

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. Gag-Pol/h is expressed from the pVR1012x/s vector backbone.

30

25

20

WO 02/32943

### **VRC4001**

pVR1012x/s SIVGag-Pol/h

Eukaryotic vector with humanized codons expressing the Gag-pol gene of SIVmac239. The Sal I –XbaI fragment of SIV gag(VRC3901) was inserted into Sal I –XbaI of SIV Pol(VRC4101) to create VRC4001.

PCT/US01/25721

## **VRC4100**

pVR1012x/s Pol/h

10

1.5

The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene using codons optimized for expression in human cells. To initiate translation at the beginning of Pol, a methionine codon was added to the 5'-end of the synthetic polymerase gene to create the Pol/h gene. Pol/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

### **VRC4101**

pVR1012x/s SIV Pol/h

20

25

The protein sequence of the pol polyprotein (amino acids 3-1017) from SIVmac239 (GenBank accession number M19921) was used to create a synthetic version of the pol gene using codons optimized for expression in human cells. To initiate translation at the beginning of Pol, a methionine codon was added to the 5'-end of the synthetic polymerase gene to create the Pol/h gene. The Protease (Pr) mutation is at pol amino acid 123 and is AAG->GGA or amino acids R->G. Reverse transcriptase(RT) mutation is at aa 352(GAC to CAT D(D to H) -RT) and Integrase mutation is at aa 788(GAC to GGC(D to A) -IN). Pol/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

30

## **VRC4200**

pVR1012x/s Gag(fs)Pol/h

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the

gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone.

## **VRC4300**

5

10

15

25

30

35

pVR1012 Gag-Pol(d delta RT delta IN)/h

Eukaryotic vector with humanized codons expressing the Gag and the frame shifted Pol genes of HIV HXB2 subtype B with deletions in Reverse transcriptase, and Integrase regions. The protein sequence of the gag polyprotein (Pr55, amino acids1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express

the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6 The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. Note: This vector is not in the pVR1012x/s backbone.

# **VRC4301**

5

10 .

15

20

25

30

35

pVR1012x/s-Gag(FS)-Pol-delta RT IN-IRES-R5 gp157-Nef

For the Gag(FS)Pol delta RT delta IN/h portion, VRC4302 pVR1012x/s Gag(fs)Pol(delta PR delta/h was used for the protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy- terminus of the gag polyprotein, pl and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid

771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gagpol amino acid 1209 and is ACT->CAT or amino acids D->A. This gene has been fused to an Internal Ribosomal Entry Site (IRES) and then fused to the R5gp157-Nef from VRC2200 pVR1012x/s R5gp157- NefDMHCDCD4/h in which the protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The envelope-Nef fusion protein expressed from pR5gp157-Nef/h contains the first 820 amino acids from the HIV envelope glycoprotein (gp157) fused to the entire mutant Nef protein. The gene for gp157 was ligated in frame with the full-length mutant Nef gene from pNefDMHCDCD4/h (VRC3600) to produce pR5gp157- NefDMHCDCD4/h. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both MHC class I and CD4 expression, point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. R5gp157-NefDMHCDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

#### **VRC4302**

pVR1012 Gag(delFS)Pol(delta PR delta RT delta IN)/h

30

35

5

10

15

20

25

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag

proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, pl and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag- pol amino acid 1209 and is ACT->CAT or amino acids D->A. Note: This vector is not in the pVR1012x/s backbone.

# **VRC4303**

5

10

15

20

35

pVR1012 SIV Gag(delFS)Pol(delta PR delta RT delta IN)/h

Eukaryotic vector with humanized codons expressing the Gag and the frame-shift-deleted Pol genes SIVmac239 with deletions in the Protease, Reverse transcriptase, and Integrase regions. The 5 of Ts from 3188 to 3192 of SIV gag-pol(VRC4001) were deleted to create VRC4303.

## 30 VRC4304

pVR1012 Gag (delFS)Pol delta PR delta RT delta IN/h

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HIV-1 C clade(GenBank accession number U52953) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The

nucleotide sequence of the synthetic gag gene shows little homology to the HIV-1 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, pl and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including pl and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag- pol amino acid 1209 and is ACT->CAT or amino acids D->A. Note: This vector is not in the pVR1012x/s backbone.

#### **VRC4305**

5

10

15

20

30

35

pVR1012 Gag-A(delFS)Pol(delta PR delta RT delta IN)/h

Eukaryotic vector with humanized codons expressing the Gag and the frame shifted Pol genes of HIV HIV-1A clade with deletions in the Protease, Reverse transcriptase, and Integrase regions. VRC4305 pVR1012 Gag(-AdelFS)Pol(delta PR delta RT delta IN)/h The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HIV-1 A clade (GenBank accession number AF004885) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HIV-1 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally

cleaved from the carboxy-terminus of the gag polyprotein, pl and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag- pol amino acid 1209 and is ACT->CAT or amino acids D->A. Note: This vector is not in the pVR1012x/s backbone.

20

25

30

35

5

10

15

# pVRC4306

pVR1012 Gag(delFS)Pol delta PR delta RT delta IN/Nef/h

The protein sequence of the Gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the Gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic Gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic Gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the Gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic Gag gene was ligated in frame with sequences encoding the Pol polyprotein. The protein sequence of the Pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the Pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the Gag-Pol

polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of Gag plus an additional 3 amino acids from Pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all Gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. The Protease (PR) mutation is at Gag-Pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at Gag-Pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at Gag- Pol amino acid 1209 and is ACT->CAT or amino acids D->A. The Nef/h gene was fused to downstream of Pol gene of Gag(delFS)Pol(delta PR delta RT delta IN). No loss or extra-amino acid was created by the fusion between Nef and Pol. The ATG of Nef was preserved. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. The nucleotide sequence Nef/h shows little homology to the viral gene, but the protein encoded is the same. Note: This vector is not in the pVR1012x/s backbone.

20

25

30

35

5

10

15

# **VRC 4308**

VRC4302-myr

pVR1012 Gag (delFS) Pol deltaPR deltaRT deltaIN deltaMyr/h

The myristoylation site was deleted from pVRC4302. The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons

optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. The myristylation site was deleted. Note:This vector is not in the pVR1012x/s backbone.

### **VRC 4309**

5

10

15

20

25

30

35

pVR1012 Gag (delFS) Pol deltaPR deltaRT deltaIN delta Myr/Nef/h

The myristoylation site was deleted from pVRC4306. The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons. optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, pl and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank

accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. The stop codon TAG was removed and synthetic B clade Nef (Genbank access number) was fusion to the 3' end of pol by PCR. Note: This vector is not in the pVR1012x/s backbone.

### **VRC 4310**

pVR1012 Nef Gag (del fs) (del Myr) Pol (delta PR delta RT delta IN)/h

15

20

25

30

35

10

5

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including pl and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag Gag(fs)Pol/h is expressed from the pVR1012x/s vector and Pol sequences.

backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. The stop codon TAG of synthetic B clade Nef (Genbank access number) gene was removed and was fused to the 5' end of gene gene by PCR. Note: This vector is not in the pVR1012x/s backbone.

## **VRC 4311**

5

15

20

25

30

35

10 pVR1012 Gag-C(delFS)Pol(delta PR delta RT delta IN) Nef/h VRC4304 + Clade C Nef

> The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HIV-1 C clade(GenBank accession number U52953) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. nucleotide sequence of the synthetic gag gene shows little homology to the HIV-1 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including pl and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN)

mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. The stop codon TAG was removed and synthetic C clade Nef (Genbank accession number:U52953) was fusion to the 3' end of pol by PCR. Note:This vector is not in the pVR1012x/s backbone.

5

10

15

20

25

30

### **VRC 4312**

Gag (del fs) (del Myr) Nef Pol ΔPRΔRTΔIN/h

VRC-Myr-gag-Dnef-Dpol

(Eukaryotic vector with humanized codons expressing the Gag, truncated Nef, and truncated Pol proteins of HIV subtype B .ns.)

This construct was derived from VRC4302. The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, pl and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding synthetic nef gene that 51 aa were deleted from 5'. 77 aa were deleted from 5' of pol polyprotein, and ligated with 3' of nef in which tag stop codon was deleted. The protein sequence of the polyprotein (amino acids 3-78 -1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including pl and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG-

>GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. Note: This vector is not in the pVR1012x/s backbone.

5

10

15

20

25

30

# **VRC 4313**

pVR1012 Gag Clade A (del fs)Pol( $\Delta$  PR  $\Delta$  RT  $\Delta$  IN)/h VRC4305+ Clade A Nef

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HIV-1 A clade (GenBank accession number AF004885) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HIV-1 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The gene was ligated in frame with sequences encoding the pol synthetic gag polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including pl and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. The stop codon TAG was removed and synthetic A clade Nef ( Genbank accession

number:AF069670) was fused to the 3' end of pol by PCR. Note: This vector is not in the pVR1012x/s backbone.

## Examples

5

10

15

20

25

30

#### Env

### **Immunogens**

Plasmids expressing the CXCR4-tropic HIV-1 HXB2 Env were made synthetically with sequences designed to disrupt viral RNA structures that limit protein expression by using codons typically found in human cells. Briefly, the synthetic Env gene of HXB2 (GenBank accession number K03455) was generated in three fragments by assembling the overlapping synthetic oligonucleotides using PCR amplification. To produce a CCR5-tropic version of the HIV-1 envelope, the region encoding amino acids 275 to 361 of HXB2 (CXCR4-tropic) gp160 was replaced with CCR5-tropic HIV-1 BaL sequence (GenBank accession number M68893), which includes the V3 loop. Glycosylation mutants were generated by site-directed mutagenesis to replace asparagine with glutamic acid residues at seventeen conserved glycosylation sites between amino acids 88 and 448. To express truncation mutant Env proteins, stop codons were introduced after positions 752, 704, 680 or 592 to produce gp150, gp145, gp140, or gp128, respectively. The Env protein was further changed by deleting amino acids 503 to 537 and 593 to 619, which removes the cleavage site sequence, the fusion domain, and a part of the spacer between the two heptad repeats. The structures of the synthetic HIV envelope genes are shown (Fig. I). The cDNAs were cloned in the expression vector pNGVL or pVR1012 under the control of the cytomegalovirus immediate-early enhancer, promoter, and first intron. The protein sequence was identical to HXB2 Env except for the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T.

# **Expression of Envelope Proteins in Transfected Cells**

293 cells (10<sup>6</sup>) were plated in 60mm dishes. Cells were transfected on the following day with 2 μg of plasmid using calcium phosphate (Graham FL, van der Eb AJ, 1973, Virology, 52:456-67). Cells were harvested 48 hours after transfection and lysed in buffer containing 50mM HEPES pH 7.0, 250mM NaCl and 0.5% NP40. The protein concentration in the lysates was determined using the Bradford reagent (Bio-Rad). Proteins (25 μg) in lysates were separated by 7.5% SDS-PAGE

and transferred to Immobilion-P membrane (Millipore, Bedford, MA). Env was detected by immunoprecipitation followed by Western blotting using polyclonal antibody against gp160 (Intracel, Rockville, MD).

# Cell Surface Expression of Envelope by FACS Analysis

293 cells were harvested 48 hours after transfection and washed twice with phosphate-buffered saline (PBS) containing 1% bovine serum albumin and incubated for 30 minutes on ice with polyclonal immunoglobulin from an HIV-1 infected patient. The secondary antibody against human IgG conjugated with FITC (Jackson Immuno Research) was added, incubated for 30 minutes on ice, washed 3 times with PBS, and analyzed by flow cytometry (FACScan). The median fluorescent intensity values were derived using Cell Quant software.

#### **DNA** Injection in Mice

5

10

15

20

25

30

Six week old, female BALB/c mice were injected intramuscularly with 100 µg of purified plasmid DNA suspended in 200 µl of normal saline. For each plasmid DNA, a group of 4 mice was injected three times at intervals of two weeks. The mice were bled two weeks after the last injection, sera collected and stored at 4°C.

# Quantitation of the Antibody Response

Immunoprecipitation and Western blotting was used to detect the antibodies that bind to native envelope proteins. Sera from immunized mice were used to immunoprecipitate gp160 from cell lysates of gp160-transfected 293 cells. Indicated dilutions were used to immunoprecipitate gp160 from the cell lysate (400 µg). Immunocomplexes were separated by 7.5% SDS-PAGE and analyzed by immunoblotting using the polyclonal antibody against gp160.

#### Analysis of CTL Response

Spleens were removed aseptically and gently homogenized to a single cell suspension, washed, and resuspended to a final concentration of  $5x10^7$  cells/ml. Cells were incubated for 7 days in presence of IL-2 (10 U/ml) and either irradiated peptide-pulsed splenocytes from naive mice or an irradiated stable cell line expressing the full length gp160 BC-env/rev, (Irwin MJ et al., 1994, J Virol, 68:5036-44). Three types of target cells were used: peptide-pulsed P815 cells (ATCC TIB64), BC10ME cells stably expressing gp160, and BC10ME cells (Collins JL, Patek PQ, Cohn M. Tumorigenicity and lysis by natural killers. J Exp Med, 1981; 153:89-106) pulsed with peptides derived from gp160 sequence. Target

cells were labeled with <sup>51</sup>Cr for 90 minutes and washed three times with RPMI-1640 with 10% FBS, 2mM glutamine, 5x10<sup>-5</sup>M β-mercaptoethanol and fungisone, 250units/ml, and resuspended in this media. Cytolytic activity was determined in triplicate samples using all different target cell dilutions in a 5-hour <sup>51</sup>Cr-release assay (Ohno T et al., 1997, Gene Ther, 4:361-6).

## Gag and Pol

# Development of Synthetic HIV-1 Gag-Pol Expression Vectors

5

10

15

20

25

30

The protein sequences of Gag (amino acids 1-432) from HXB2 (GenBank accession number K03455) and Pol (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) were reverse translated by the GCG Package (Genetic Computer Group, Inc., Madison, WI) using codons expected for human cells. A 226bp fragment spanning the frame shift site and the overlapping region of the two reading frames from NL4-3 were retained to allow expression of Gag and Gag-Pol precursor polyproteins in the same construct. 86 oligonucleotides covering 4325 DNA base pairs with 5' SalI and 3' EcoRI sites were purchased from GIBCO Life Technologies. Each of the oligonucleotides was 75 base-pairs with 25 nt of overlap. The codon optimized Gag-Pol gene (hGag-Pol) was assembled by PCR with Pwo (Boehringer Mannheim) and Turbo Pfu (Stratagene) high fidelity DNA polymerase. The PCR conditions were optimized with a PCR optimization kit (Stratagene) on a gradient Robocycler (Stratagene). Full-length synthetic Gag-Pol gene was cloned into the Sal I and blunted Bgl II site of the mammalian expression vector pNGVL-3 (Xu, Ling et al., 1998, Nature Medicine, 4:37-42) and confirmed by DNA sequencing. Three additional constructs were derived from the hGag-Pol gene. 5 Thymidines (Ts) in the frame shift site (FS) of the hGag-Pol gene were deleted (ΔFS) and the protease was inactivated by replacing AGG in protease to GGC (R42G) to create hGag-Pol\DeltaFS\DeltaPr (Hung, M, 1998, J Virol, 72:4819-4824; Loeb, DD et al., 1989, Nature 340:397-400). 432 amino acids of the NH2- terminal of hGag-Pol gene were deleted and an ATG start codon was added to create the hPol gene. 925 amino acids of the COOH-terminal hGag-Pol were deleted to create the hGag gene. hGagPol\DeltaFS\DeltaPr, hPol and hGag genes were expressed in the pNGVL-3 plasmid, derived by insertion of a polylinker into pVR1012. The plasmid expressing viral Gag-Pol, pCMVΔ8.2 was a kind gift (Naldini, L et al., 1996, Science, 272:263-267).

## Transfert Transfection and Analysis of Expression

5

10

15

20

25

30

293T cells were maintained in Dulbecco's modified Eagle medium (DMEM; GIBCO-BRL), supplemented with 10% fetal bovine serum (FBS). Plasmid DNAs were purified with double cesium chloride sedimentation gradients. Approximately 3×10<sup>6</sup> 293T cells were placed in a 10-cm dish one day before transfection. 10 μg of pCMVdR8.2 plasmid (containing the viral gag-pol gene), or 5 µg of pVR1012s (containing the codon altered genes), were used to transfect 293T cells, using the calcium phosphate method (Chen, C and H Okayama, 1987, Mol. Cell. Biol, 7:2745-2752). Three days after transfection, cell lysates were prepared with RIPA buffer (Boehringer Mannheim, Indianapolis, IN) and separated by 4-15% gradient SDSpolyacrylamide gel electrophoresis (PAGE), then transferred onto an Immobilon P membrane (Millipore). Membranes were then incubated with anti HIV-1-IgG (AIDS Research and Reference Reagent Program), monoclonal anti-p24 (ICN), or rabbit anti-RT (Intracel, Rockville, MD). Bands were visualized using the ECL Western blotting detection reagent (Amersham Pharmacia Biotech, Piscataway, NJ), as described by the manufacturer. Expression levels were determined using a phosphorimager.

# Generation of Stably Transfected Cell Lines

hGag and hPol genes were individually subcloned into the Xho I and EcoRI sites of a retroviral vector, pPGS-CITE-Neo. Three plasmid systems were used to produce recombinant retroviruses containing the hGag or hPol genes (Yang, Z et al., 1998, Science, 279:1034-1037). 48 hours after transfection, the supernatants were collected to transduce CT26 and BC10ME (Collins, JL, 1981, J Exp. Med, 153:89-106) which are syngeneic to Balb/C mice, and selected in 0.8 mg/ml of G418 two days after infection. The positive clones were screened and confirmed by Western blotting and maintained in 10% FCS supplemented RPMI (GIBCO-BRL) with 0.5mg/ml G418.

## **DNA Vaccination of Mice**

Female BALB/C mice, 6-8 weeks old, were used for immunogenicity studies. For DNA vaccination, mice were immunized with 100  $\mu$ l (0.5  $\mu$ g/ml DNA, 0.9% NaCl) in the quadriceps muscle of each hind leg every two weeks for a total of 4 injections.

### CTL Assay

5

1.5

20

30

Animals were sacrificed after the last immunization and spleens removed from both naïve and immunized mice using aseptic techniques. Splenic lymphocytes were harvested and the chromium release CTL assay was performed in triplicate as previously described (Ohno, T. et al., 1997, Gene Ther, 4:361-366). The peptides used for sensitizing cells are as follows: Two peptide mixtures from the Gag protein P17(88-115) and p24(62-76). Seven peptide mixtures from the Pol protein: 1) P66(175-189), 2) P66(179-193), 3) P66(183-197), 4) P66(187-201), 5) P66(223-237), 6) P66(227-241), 7) P66(367-381).

# 10 Measurement of Antibody Responses

Anti-p24 ELISA assays

The anti-p24 ELISA assay was performed in Immunlon ninety-six-well plates (Dynet Technologies Inc., Chantilly, VA). The plates were coated with 50 μl of purified recombinant HIV-1IIIB p24 antigen (Intracel) at a concentration of 2 μg/ml in PBS buffer, pH 7.4 (GIBCO) with 0.05% sodium azide. Plates were washed 3x in PBS containing 3% BSA and 0.05% Tween20 (blocking buffer) and incubated for 2 hours. Mouse sera were serially diluted from 1:100 to 1:12,800 in blocking buffer, added to the p24-coated plates and incubated overnight at 4°C. Plates were then washed four times with PBS (0.05% Tween 20), and incubated with goat anti-mouse IgG (1:10,000 dilution, Sigma) for 2 hours at room temperature. Plates were washed four times, and then pNPP alkaline phosphatase substrate (75 μl; Sigma, St. Louis, MO) was added to each well. The reaction was stopped after one hour by addition of 0.5 N NaOH (25 μl). The plates were read on an ELISA reader at 405 nm, and titers were calculated at a cutoff optical density of 0.4.

## 25 HIV-1 immunoblotting

The strips containing HIV-1 proteins (Immunectics Inc., Cambridge, MA) were incubated with pooled mouse sera at a dilution of 1:25. Purified human anti-HIV IgG (AIDS Research and Reference Reagent Program, Rockville, MD) was used as a positive control. Bands were visualized using the ECL Western blotting detection reagent (Amersham Pharmacia Biotech, Piscataway, NJ).

# Immunoprecipitation and Western blotting

Three days after transfection, hPol gene-transfected 293T cell lysates were prepared with RIPA buffer. The pooled mouse sera were diluted with immunoprecipitation (IP) buffer (100mM KCl, 2.5mM MgCl<sub>2</sub>, 20mM HEPES, pH

7.9, 0.1% NP-40, 1mM DTT and proteinase inhibitors). After adding 10 µg of the cell lysate containing the HIV-1 Pol protein, the reactions were incubated overnight on a rotator at 4°C. 250 µl of Protein G and A Sepharose beads (10%V/V in IP buffer) were then added, and the reactions were incubated for 2 hours on a rotator at 4°C. The beads were washed four times with IP buffer, resuspended in 30 µl of IX sample buffer and loaded onto SDS-PAGE. After transfer onto an Immobilon P membrane (Millipore), membranes were incubated with anti HIV-1-IgG (AIDS Research and Reference Reagent Program). Bands were visualized using the ECL Western blotting detection reagent (Amersham Pharmacia Biotech).

10

15

5

# Prime/Boost Vaccination Strategy

Groups of guinea pigs (4/group) were immunized (primed) with the vectors listed below, 3 times at 2 week intervals. Two weeks after the third DNA immunization, blood was collected for immune analysis. 2-3 days after the blood collection, the animals received a boost with a corresponding adenoviral construct, AdApt (Sullivan, N. et al., 2000, Nature, 408:605-608). Blood was again collected 2 weeks after the adenoviral boost. Blood was tested for neutralizing antibodies and (using an ELISA) the presence of antibody against the envelope. The groups are summarized below.

|    | Prime                   | Adenoviral Boost       |
|----|-------------------------|------------------------|
| 20 | gp140delCFI (R5)        | Adv gp140delCFI(R5)    |
|    | gp145delCFI (R5)        | Adv gp145delCFI(R5)    |
|    | gp145delCFI (R5)        | Adv gp140delCFI(R5)    |
|    | gp140delCFI (89.6P)     | Adv gp140delCFI(89.6P) |
|    | gp145delCFI (89.6P)     | Adv gp145delCFI(89.6P) |
| 25 | gp145delCFI (89.6P)     | Adv gp140delCFI(89.6P) |
|    | Adv gp140delCFI (R5)    | Adv gp140delCFI(R5)    |
|    | Adv gp145delCFI (R5)    | Adv gp145delCFI(R5)    |
|    | Adv gp145delCFI (R5)    | Adv gp140delCFI(R5)    |
|    | Adv gp140delCFI (89.6P) | Adv gp140delCFI(89.6P) |
| 30 | Adv gp145delCFI (89.6P) | Adv gp145delCFI(89.6P) |
|    | Adv gp145delCFI (89.6P) | Adv gp140delCFI(89.6P) |

Results

The level of antibody in the sera after adenoviral boost was 8- to10-fold higher than the level of antibody present after 3 injections of DNA alone, in almost all of the cases. Neutralizing antibodies were detected in the sera of animals immunized with either gp145delCFI DNA or gp140delCFI DNA followed by Advgp140delCFI.

Neutralization was also observed in animals immunized with gp145delCFI DNA followed by Adv gp145delCFI, and in animals both primed and boosted with adenovirus.

10

5

Optimal neutralization was seen when animals were primed with gp145delCFI DNA followed by adenovirus expressing gp140delCFI in both clade B (R5-Bal strain) and 89.6P (Clade B, dual tropic strain).

All neutralizing activity was able to be blocked with the V3 peptide of the corresponding strains.

15

20

\*\*\*\*

While the present invention has been described in some detail for purposes of clarity and understanding, one skilled in the art will appreciate that various changes in form and detail can be made without departing from the true scope of the invention. All figures, tables, and appendices, as well as patents, applications and publications referred to above are hereby incorporated by reference.

### WHAT IS CLAIMED IS:

5

10

15

20

25

. 30

35

1. A nucleic acid molecule comprising a polynucleotide encoding a modified glycosylation or  $\Delta$ CFI HIV Env, or combination thereof, encoding a plasmid taken from Table 1 (or its modified HIV Env insert), or encoding an analog of said plasmid or insert having at least 95% identity thereto.

- 2. The nucleic acid molecule of Claim 1 which is plasmid 2100 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 3. The nucleic acid molecule of Claim 1 which is plasmid 2200 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 4. The nucleic acid molecule of Claim 1 which is plasmid 2300 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 5. The nucleic acid molecule of Claim 1 which is plasmid 2302 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 6. The nucleic acid molecule of Claim 1 which is plasmid 2400 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 7. The nucleic acid molecule of Claim 1 which is plasmid 2700 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 8. The nucleic acid molecule of Claim 1 which is plasmid 2701 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 9. The nucleic acid molecule of Claim 1 which is plasmid 2702 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 10. The nucleic acid molecule of Claim 1 which is plasmid 2706 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 11. The nucleic acid molecule of Claim 1 which is plasmid 2707 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

10

15

20

25

30

12. The nucleic acid molecule of Claim 1 which is plasmid 2800 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 13. The nucleic acid molecule of Claim 1 which is plasmid 2801 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 14. The nucleic acid molecule of Claim 1 which is plasmid 2804 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 15. The nucleic acid molecule of Claim 1 which is plasmid 2805 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 16. The nucleic acid molecule of Claim 1 which is plasmid 2810 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 17. The nucleic acid molecule of Claim 1 which is plasmid 2811 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 18. The nucleic acid molecule of Claim 1 which is plasmid 2812 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 19. The nucleic acid molecule of Claim 1 which is plasmid 2813 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 20. The nucleic acid molecule of Claim 1 which is plasmid 2814 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 21. The nucleic acid molecule of Claim 1 which is plasmid 2820 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 22. The nucleic acid molecule of Claim 1 which is plasmid 2821 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

15

20

25

30

23. The nucleic acid molecule of Claim 1 which is plasmid 2822 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 24. The nucleic acid molecule of Claim 1 which is plasmid 2823 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 25. The nucleic acid molecule of Claim 1 which is plasmid 2824 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 10 26. The nucleic acid molecule of Claim 1 which is plasmid 2825 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 27. The nucleic acid molecule of Claim 1 which is plasmid 2826 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 28. The nucleic acid molecule of Claim 1 which is plasmid 2827 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 29. The nucleic acid molecule of Claim 1 which is plasmid 2828 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 30. The nucleic acid molecule of Claim 1 which is plasmid 2829 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 31. The nucleic acid molecule of Claim 1 which is plasmid 2830 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 32. The nucleic acid molecule of Claim 1 which is plasmid 2831 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 33. The nucleic acid molecule of Claim 1 which is plasmid 2832 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

10

15

20

25

30

34. The nucleic acid molecule of Claim 1 which is plasmid 2833 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 35. The nucleic acid molecule of Claim 1 which is plasmid 2834 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 36. The nucleic acid molecule of Claim 1 which is plasmid 2835 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 37. The nucleic acid molecule of Claim 1 which is plasmid 2836 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 38. The nucleic acid molecule of Claim 1 which is plasmid 2837 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 39. The nucleic acid molecule of Claim 1 which is plasmid 2838 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 40. The nucleic acid molecule of Claim 1 which is plasmid 2839 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 41. The nucleic acid molecule of Claim 1 which is plasmid 2840 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 42. The nucleic acid molecule of Claim 1 which is plasmid 2841 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
    - 43. The nucleic acid molecule of Claim 1 which is plasmid 2842 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
    - 44. The nucleic acid molecule of Claim 1 which is plasmid 2843 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

15

20

25

30

45. The nucleic acid molecule of Claim 1 which is plasmid 2844 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 46. The nucleic acid molecule of Claim 1 which is plasmid 2845 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 47. The nucleic acid molecule of Claim.1 which is plasmid 2846 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 10 48. The nucleic acid molecule of Claim 1 which is plasmid 2847 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 49. The nucleic acid molecule of Claim 1 which is plasmid 2848 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 50. The nucleic acid molecule of Claim 1 which is plasmid 2849 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 51. The nucleic acid molecule of Claim 1 which is plasmid 2850 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 52. The nucleic acid molecule of Claim 1 which is plasmid 2851 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 53. The nucleic acid molecule of Claim 1 which is plasmid 2852 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 54. The nucleic acid molecule of Claim 1 which is plasmid 2853 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 55. The nucleic acid molecule of Claim 1 which is plasmid 2854 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

15

20

25

30

- 56. The nucleic acid molecule of Claim 1 which is plasmid 2860 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 57. The nucleic acid molecule of Claim 1 which is plasmid 2861 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 58. The nucleic acid molecule of Claim 1 which is plasmid 2862 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- The nucleic acid molecule of Claim 1 which is plasmid 2863 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 60. The nucleic acid molecule of Claim 1 which is plasmid 2864 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 61. The nucleic acid molecule of Claim 1 which is plasmid 2865 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 62. The nucleic acid molecule of Claim 1 which is plasmid 2866 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 63. The nucleic acid molecule of Claim 1 which is plasmid 2867 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 64. The nucleic acid molecule of Claim 1 which is plasmid 2868 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
    - 65. The nucleic acid molecule of Claim 1 which is plasmid 2869 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
    - 66. The nucleic acid molecule of Claim 1 which is plasmid 2870 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

1,5

20

25

30

- 67. The nucleic acid molecule of Claim 1 which is plasmid 2871 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 68. The nucleic acid molecule of Claim 1 which is plasmid 2872 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 69. The nucleic acid molecule of Claim 1 which is plasmid 2873 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 10 70. The nucleic acid molecule of Claim 1 which is plasmid 2874 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 71. The nucleic acid molecule of Claim 1 which is plasmid 2900 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 72. The nucleic acid molecule of Claim 1 which is plasmid 3000 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 73. The nucleic acid molecule of Claim 1 which is plasmid 3200 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 74. The nucleic acid molecule of Claim 1 which is plasmid 3201 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 75. The nucleic acid molecule of Claim 1 which is plasmid 3202 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 76. The nucleic acid molecule of Claim 1 which is plasmid 3203 (or its modified HTV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 77. The nucleic acid molecule of Claim 1 which is plasmid 3300 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

10

15

20

25

30

78. The nucleic acid molecule of Claim 1 which is plasmid 3301 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 79. The nucleic acid molecule of Claim 1 which is plasmid 3400 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 80. The nucleic acid molecule of Claim 1 which is plasmid 3401 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 81. The nucleic acid molecule of Claim 1 which is plasmid 3500 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 82. The nucleic acid molecule of Claim 1 which is plasmid 3600 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 83. The nucleic acid molecule of Claim 1 which is plasmid 3700 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 84. The nucleic acid molecule of Claim 1 which is plasmid 3800 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 85. The nucleic acid molecule of Claim 1 which is plasmid 5200 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 86. The nucleic acid molecule of Claim 1 which is plasmid 5201 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 87. The nucleic acid molecule of Claim 1 which is plasmid 5202 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 88. The nucleic acid molecule of Claim 1 which is plasmid 5203 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

10

15

20

25

30

89. The nucleic acid molecule of Claim 1 which is plasmid 5300 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 90. The nucleic acid molecule of Claim 1 which is plasmid 5301 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 91. The nucleic acid molecule of Claim 1 which is plasmid 5303 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 92. The nucleic acid molecule of Claim 1 which is plasmid 5304 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 93. The nucleic acid molecule of Claim 1 which is plasmid 5305 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 94. The nucleic acid molecule of Claim 1 which is plasmid 5306 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 95. The nucleic acid molecule of Claim 1 which is plasmid 5307 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 96. The nucleic acid molecule of Claim 1 which is plasmid 5308 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 97. The nucleic acid molecule of Claim 1 which is plasmid 5309 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 98. The nucleic acid molecule of Claim 1 which is plasmid 5350 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 99. The nucleic acid molecule of Claim I which is plasmid 5351 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

10

15

20

25

30

100. The nucleic acid molecule of Claim 1 which is plasmid 5352 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 101. The nucleic acid molecule of Claim 1 which is plasmid 5353 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 102. The nucleic acid molecule of Claim 1 which is plasmid 5354 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 103. The nucleic acid molecule of Claim 1 which is plasmid 5355 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 104. The nucleic acid molecule of Claim 1 which is plasmid 5356 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 105. The nucleic acid molecule of Claim 1 which is plasmid 5357 (or its modified HTV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 106. The nucleic acid molecule of Claim 1 which is plasmid 5358 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 107. The nucleic acid molecule of Claim 1 which is plasmid 5359 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 108. The nucleic acid molecule of Claim 1 which is plasmid 5360 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 109. The nucleic acid molecule of Claim 1 which is plasmid 5361 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 110. The nucleic acid molecule of Claim 1 which is plasmid 5362 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

10

15

20

25

30

111. The nucleic acid molecule of Claim 1 which is plasmid 5363 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 112. The nucleic acid molecule of Claim 1 which is plasmid 5364 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 113. The nucleic acid molecule of Claim 1 which is plasmid 5365 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 114. The nucleic acid molecule of Claim 1 which is plasmid 5366 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 115. The nucleic acid molecule of Claim 1 which is plasmid 5367 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 116. The nucleic acid molecule of Claim 1 which is plasmid 5368 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 117. The nucleic acid molecule of Claim 1 which is plasmid 5369 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 118. The nucleic acid molecule of Claim 1 which is plasmid 5370 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 119. The nucleic acid molecule of Claim 1 which is plasmid 5371 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 120. The nucleic acid molecule of Claim 1 which is plasmid 5372 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 121. The nucleic acid molecule of Claim 1 which is plasmid 5373 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

10

15

20

25

30

122. The nucleic acid molecule of Claim 1 which is plasmid 5374 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 123. The nucleic acid molecule of Claim 1 which is plasmid 5375 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 124. The nucleic acid molecule of Claim 1 which is plasmid 5376 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 125. The nucleic acid molecule of Claim 1 which is plasmid 5377 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 126. The nucleic acid molecule of Claim 1 which is plasmid 5378 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 127. The nucleic acid molecule of Claim 1 which is plasmid 5379 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 128. The nucleic acid molecule of Claim 1 which is plasmid 5500 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 129. The nucleic acid molecule of Claim 1 which is plasmid 5501 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 130. The nucleic acid molecule of Claim 1 which is plasmid 5502 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 131. The nucleic acid molecule of Claim 1 which is plasmid 5503 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 132. The nucleic acid molecule of Claim 1 which is plasmid 5504 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

10

15

20

25

30

133. The nucleic acid molecule of Claim 1 which is plasmid 5505 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 134. The nucleic acid molecule of Claim 1 which is plasmid 5506 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 135. The nucleic acid molecule of Claim 1 which is plasmid 5507 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 136. The nucleic acid molecule of Claim 1 which is plasmid 5508 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 137. The nucleic acid molecule of Claim 1 which is plasmid 5509 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 138. The nucleic acid molecule of Claim 1 which is plasmid 5510 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 139. The nucleic acid molecule of Claim 1 which is plasmid 5511 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 140. The nucleic acid molecule of Claim 1 which is plasmid 5512 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 141. The nucleic acid molecule of Claim 1 which is plasmid 5513 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 142. The nucleic acid molecule of Claim 1 which is plasmid 5514 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 143. The nucleic acid molecule of Claim 1 which is plasmid 5515 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

15

20

25

30

144. The nucleic acid molecule of Claim 1 which is plasmid 5516 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 145. The nucleic acid molecule of Claim 1 which is plasmid 5517 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 146. The nucleic acid molecule of Claim 1 which is plasmid 5518 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 10 147. The nucleic acid molecule of Claim 1 which is plasmid 5519 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 148. The nucleic acid molecule of Claim 1 which is plasmid 5520 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 149. The nucleic acid molecule of Claim 1 which is plasmid 5521 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 150. The nucleic acid molecule of Claim 1 which is plasmid 5522 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 151. The nucleic acid molecule of Claim 1 which is plasmid 5523 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 152. The nucleic acid molecule of Claim 1 which is plasmid 5524 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 153. The nucleic acid molecule of Claim 1 which is plasmid 5525 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
  - 154. The nucleic acid molecule of Claim 1 which is plasmid 5526 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

155. The nucleic acid molecule of Claim 1 which is plasmid 5527 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 156. The nucleic acid molecule of Claim 1 which is plasmid 5528 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 157. The nucleic acid molecule of Claim 1 which is plasmid 5529 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 158. A nucleic acid molecule comprising a polynucleotide encoding in a continuous open reading frame a modified HIV Gag-Pol or Gag-Pol-Nef fusion

continuous open reading frame a modified HIV Gag-Pol or Gag-Pol-Nef fusion protein, encoding a plasmid taken from Table 1 (or its modified HIV Gag, Pol, or Nef insert), or encoding an analog of said plasmid or insert having at least 95%

identity thereto.

5

15

20

25

35

159. The nucleic acid molecule of Claim 158 which is plasmid 3900 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

160. The nucleic acid molecule of Claim 158 which is plasmid 3901 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

161. The nucleic acid molecule of Claim 158 which is plasmid 4000 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having

at least 95% identity thereto.

162. The nucleic acid molecule of Claim 158 which is plasmid 4001 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

163. The nucleic acid molecule of Claim 158 which is plasmid 4100 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 164. The

164. The nucleic acid molecule of Claim 158 which is plasmid 4101 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

165. The nucleic acid molecule of Claim 158 which is plasmid 4200 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

.

5

- 10

15

20

25

30

166. The nucleic acid molecule of Claim 158 which is plasmid 4300 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 167. The nucleic acid molecule of Claim 158 which is plasmid 4301 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 168. The nucleic acid molecule of Claim 158 which is plasmid 4302 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 169. The nucleic acid molecule of Claim 158 which is plasmid 4303 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 170. The nucleic acid molecule of Claim 158 which is plasmid 4304 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 171. The nucleic acid molecule of Claim 158 which is plasmid 4305 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 172. The nucleic acid molecule of Claim 158 which is plasmid 4306 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 173. The nucleic acid molecule of Claim 158 which is plasmid 4308 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 174. The nucleic acid molecule of Claim 158 which is plasmid 4309 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 175. The nucleic acid molecule of Claim 158 which is plasmid 4310 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 176. The nucleic acid molecule of Claim 158 which is plasmid 4311 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5

10

15

20

25

35

177. The nucleic acid molecule of Claim 158 which is plasmid 4312 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

- 178. The nucleic acid molecule of Claim 158 which is plasmid 4313 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.
- 179. A modified HIV Env, Gag, Pol, or Nef encoded by the nucleic acid molecule of any of Claims 1-178.
- 180. The nucleic acid molecule of any of Claims 1-178 further comprising a backbone, wherein said backbone is a plasmid vector.
  - 181. The nucleic acid molecule of any of Claims 1-178 further comprising a backbone, wherein said backbone is an adenoviral vector.
- 182. A nucleic acid molecule comprising a polynucleotide encoding a modified glycosylation or ΔCFI HIV Env, or combination thereof, encoding a modified HIV Env insert taken from a plasmid taken from Table 1, or encoding an analog of said insert having at least 95% identity thereto, and further encoding in a continuous open reading frame a modified HIV Gag-Pol or Gag-Pol-Nef fusion protein, encoding a modified HIV Gag, Pol, or Nef insert taken from a plasmid taken from Table 1, or encoding an analog of said insert having at least 95% identity thereto.
- 183. A pharmaceutical composition comprising at least one nucleic acid molecule of any of Claims 1-178 and a pharmaceutically acceptable solution in a therapeutically effective dose
- 184. A vaccine composition comprising at least one nucleic acid molecule of any of Claims 1-178 and a pharmaceutically acceptable solution in a prophylactically effective dose.
- 185. A pharmaceutical composition comprising at least one modified HIV Env, Gag, Pol, or Nef of Claim 179 and a pharmaceutically acceptable solution in a therapeutically effective dose.
- 30 186. A vaccine composition comprising at least one modified HIV Env, Gag, Pol, or Nef of Claim 179 and a pharmaceutically acceptable solution in a prophylactically effective dose.
  - 187. The composition of any of Claims 183-186 additionally comprising an adjuvant.
  - 188. The composition of Claim 187 wherein said adjuvant is a cytokine.

5

10

15

20

30

189. The composition of Claim 183 or 184 which combines plasmid 4306 (gag-pol-nef) and plasmid 2827 (env).

- 190. The composition of Claim 183 or 184 which combines plasmid 4306 (gag-pol-nef Clade B), plasmid 2827 (env Clade B), plasmid 4311 (gag-pol-nef Clade C), and plasmid 5509 (env Clade C).
- 191. A method comprising the step of administering by genetic immunization a nucleic acid molecule of any of Claims 1-178 to a host to ameliorate the symptoms of AIDS or HIV infection.
- 192. A method comprising the step of administering by genetic immunization a nucleic acid molecule of any of Claims 1-178 to a host to generate an antibody or CTL response against native HIV Env, Gag, Pol, or Nef.
- 193. A method comprising the step of administering by protein immunization a modified HIV Env, Gag, Pol, or Nef of Claim 179 to a host to ameliorate the symptoms of AIDS or HIV infection.
- 194. A method comprising the step of administering by protein immunization a modified HIV Env, Gag, Pol, or Nef of Claim 179 to a host to generate an antibody or CTL response against native HIV Env, Gag, Pol, or Nef.
  - 195. The method of any of Claims 191-194 wherein said step is a primary immunization.
    - 196. The method of any of Claims 191-194 wherein said step is a boost.
  - 197. The method of Claims 191 or 192 wherein said step is a primary immunization with a plasmid vector.
  - 198. The method of Claims 191 or 192 wherein said step is a boost with an adenoviral vector.
- 25 199. The method of Claims 191 or 192 wherein said step is a primary immunization with a plasmid vector expressing gp145delCFI and further comprising a step which is a boost with an adenoviral vector expressing gp140delCFI.
  - 200. A nucleic acid molecule of any of Claims 1-178 for use by genetic immunization in amelioration of the symptoms of AIDS or HIV infection.
  - 201. A nucleic acid molecule of any of Claims 1-178 for use by genetic immunization in generation of an antibody or CTL response against native HIV Env, Gag, Pol, or Nef.
  - 202. A modified HIV Env, Gag, Pol, or Nef of Claim 179 for use by protein immunization in amelioration of the symptoms of AIDS or HIV infection.

5

10

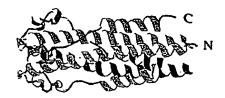
15

20

203. A modified HIV Env, Gag, Pol, or Nef of Claim 179 for use by protein immunization in generation of an antibody or CTL response against native HIV Env, Gag, Pol, or Nef.

- 204. A nucleic acid molecule comprising a polynucleotide encoding a modified  $\Delta$ CFI HIV Env or encoding in a continuous open reading frame a modified HIV Gag-Pol or Gag-Pol-Nef fusion protein, or a combination thereof, applied to any HIV clade or strain.
- 205. A nucleic acid molecule comprising a polynucleotide encoding a modified ΔCFI HIV Env or encoding in a continuous open reading frame a modified HIV Gag-Pol or Gag-Pol-Nef fusion protein, or a combination thereof, applied to related proteins of other viruses.
- 206. A nucleic acid molecule comprising a polynucleotide encoding a modified viral membrane fusion protein with deletions in sufficient domains to expose a core protein for optimal antigen presentation, wherein the viral membrane fusion protein is defined by a central triple stranded coiled coil followed by a disulfide-bonded loop that reverses the chain direction and connects to an  $\alpha$  helix packed antiparallel to the core helices.
- 207. A nucleic acid molecule comprising a polynucleotide encoding in a continuous open reading frame a modified viral protein, wherein the viral protein is transcribed in native form by translational frame shifting in a non-equivalent ratio or through a separate open reading frame.

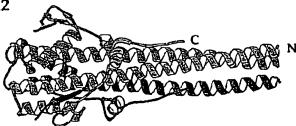
A Ebola GP2



B MoMuLv Mo-55



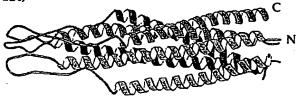
C low-pH-treated influenza HA2



D protease resistant core of HIV gp41



E SIV gp41 (NMR)



F SNAREs



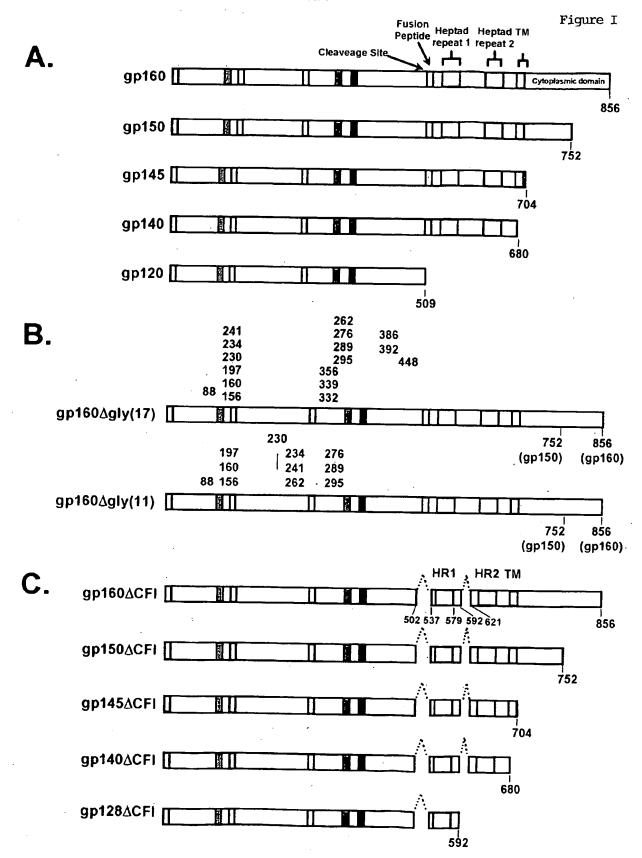
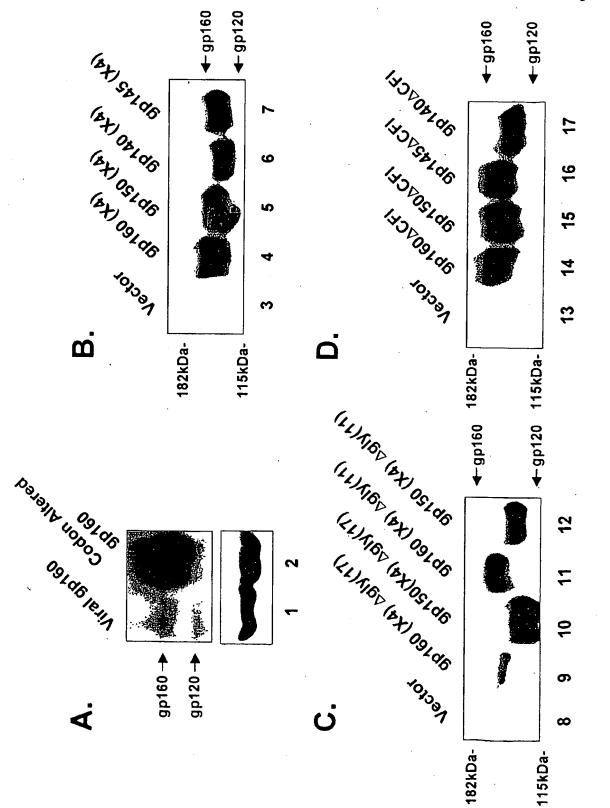


Figure II



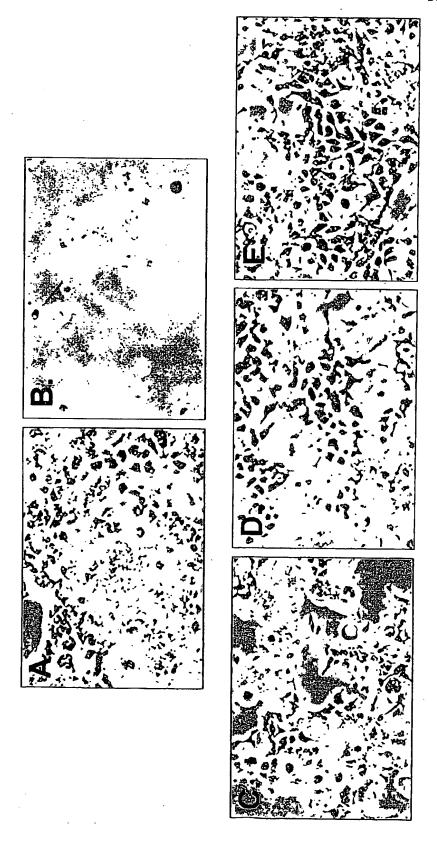
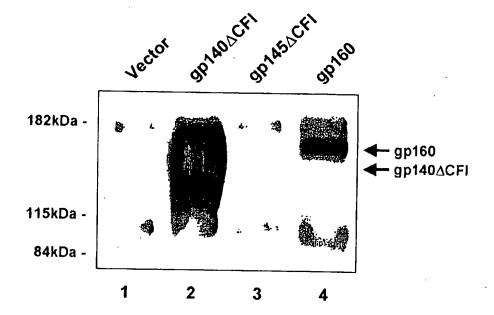
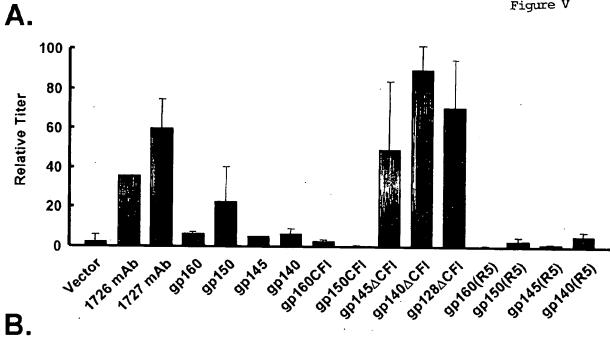
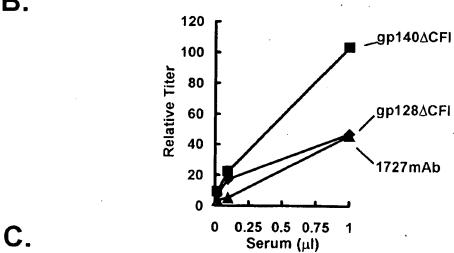


Figure IV









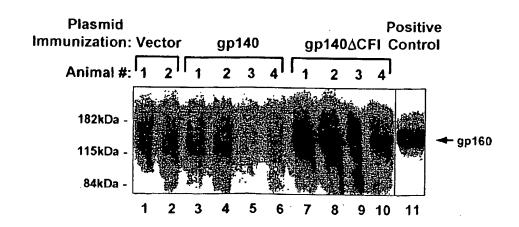
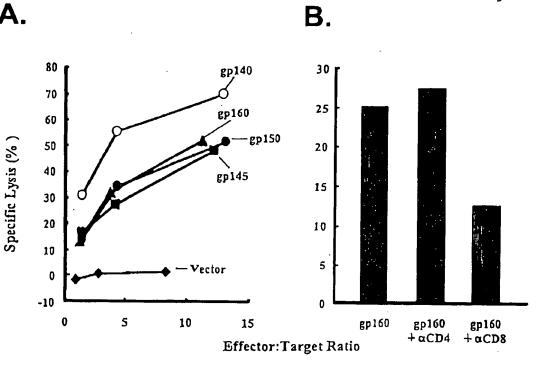


Figure VI



C. D.

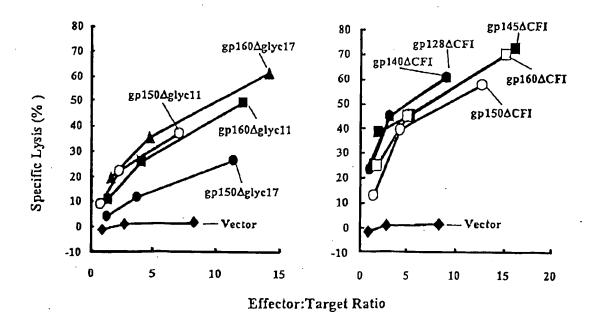
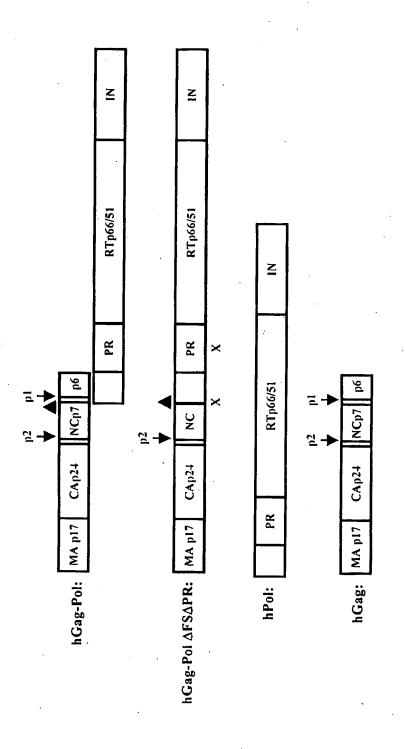


Figure VII



▲ Frame shift site (FS)

Figure VIII

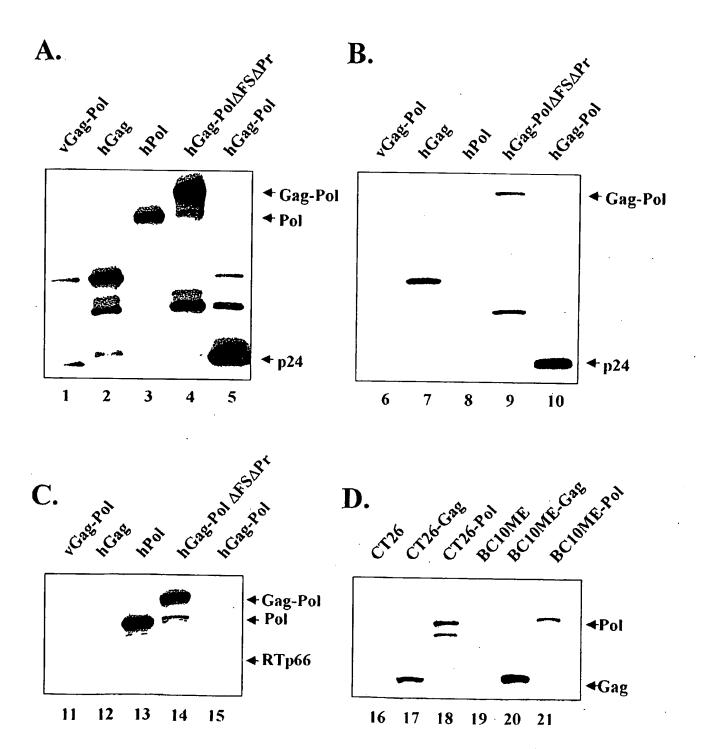
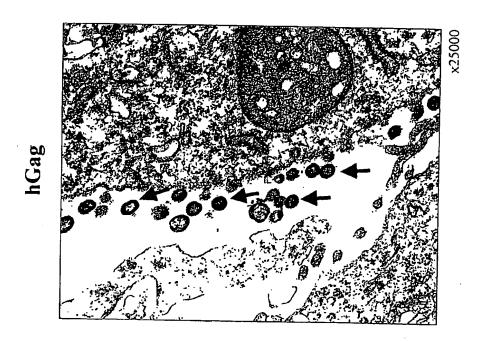


Figure IX



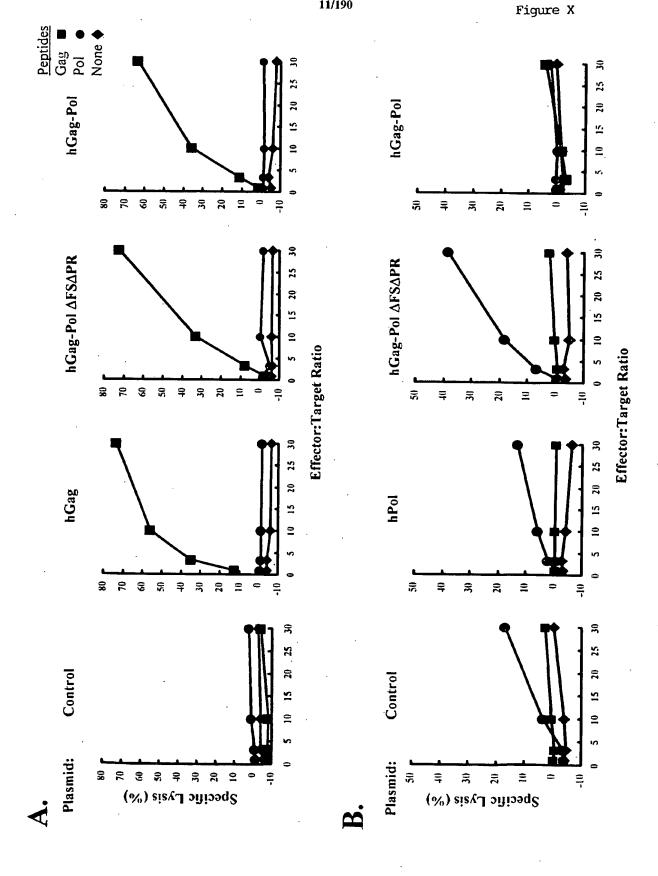


Figure X

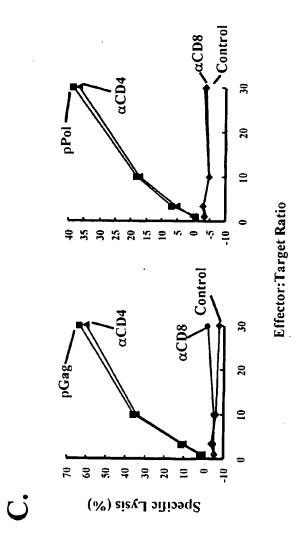
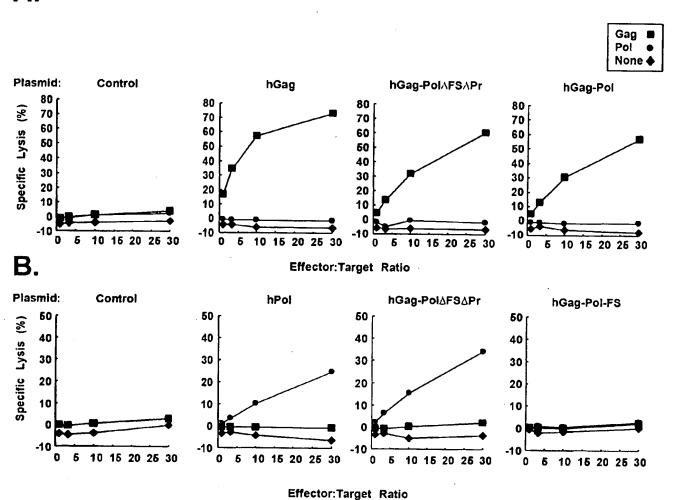
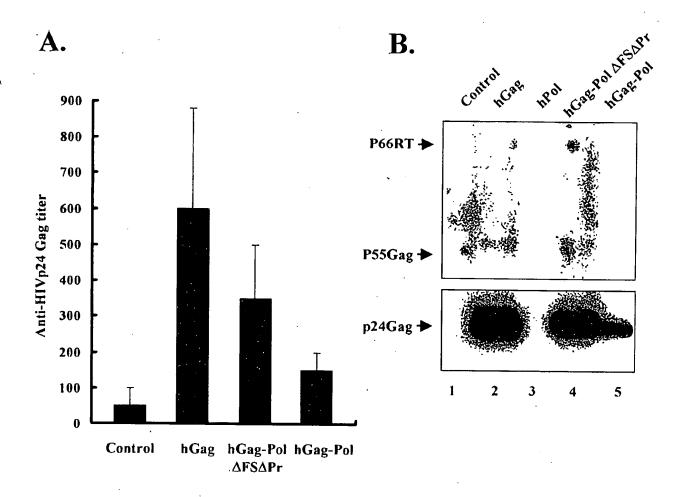
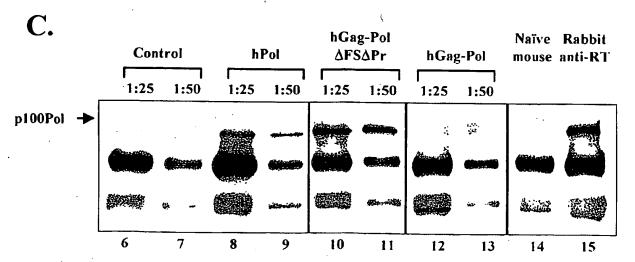


Figure XI

A.







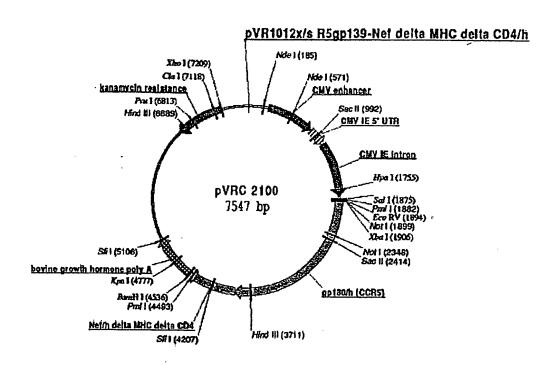


Figure 1

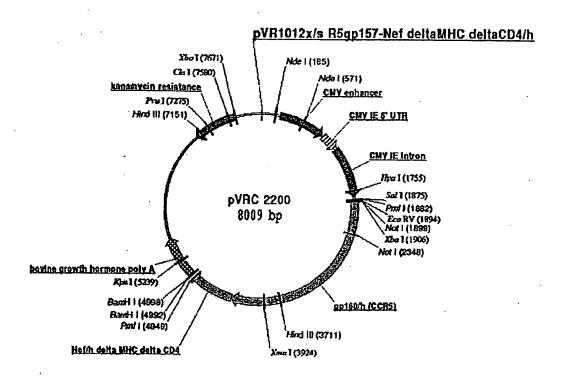


Figure 2

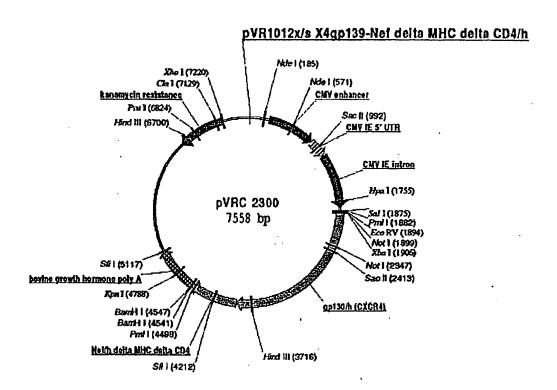


Figure 3

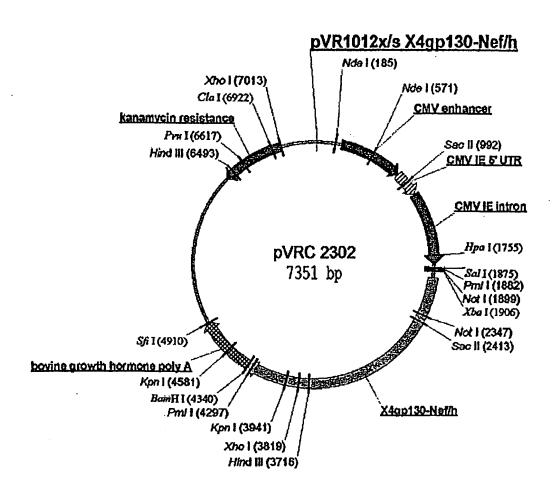


Figure 4

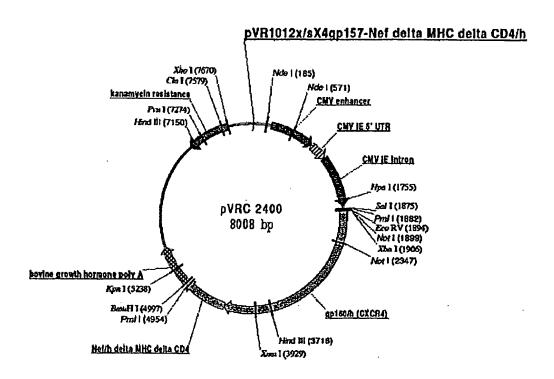


Figure 5

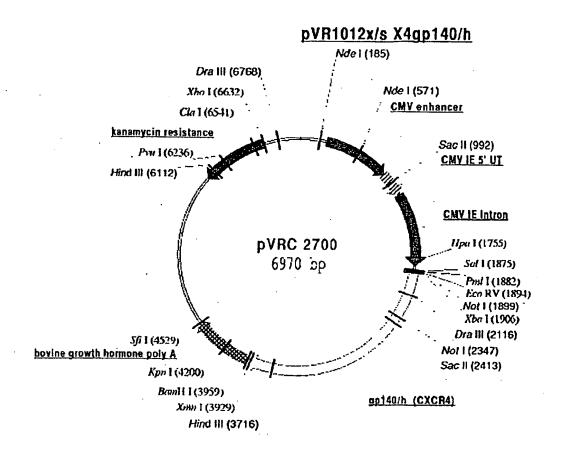


Figure 6

#### pVR1012x/s X4gp140(del F/CL del H IS)/h

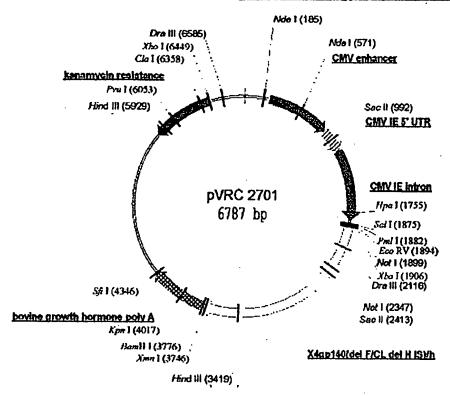


Figure 7

## pVR1012x/s X4qp128(del F/CL)/h

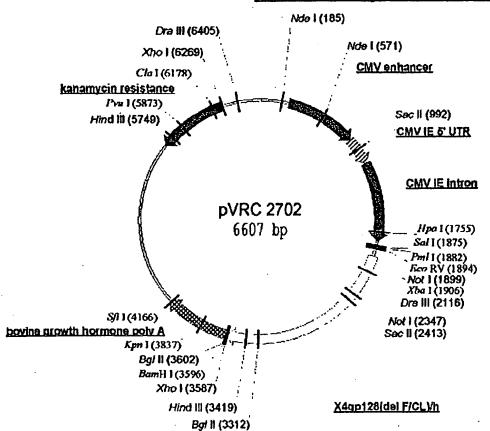


Figure 8

# pVR1012x/s X4qp145/h

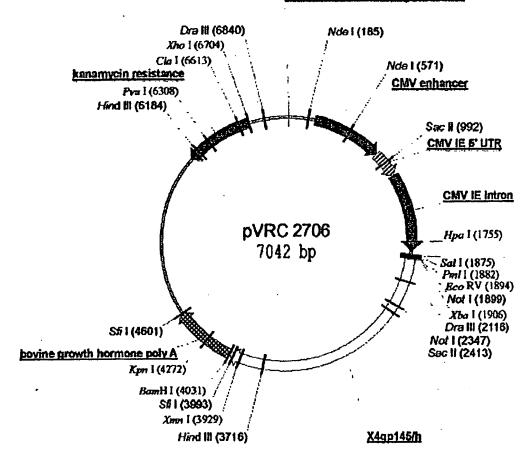


Figure 9

## pVR1012x/s X4qp145(del F/CL del H IS)/h

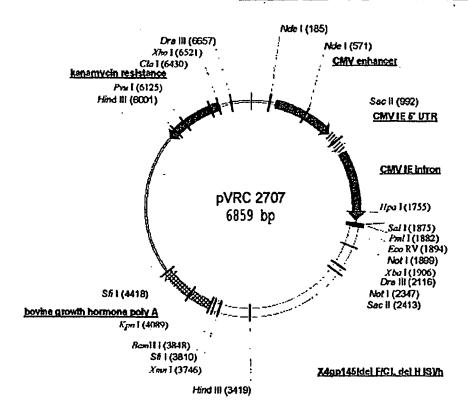


Figure 10

# pVR1012x/s R5gp140/h

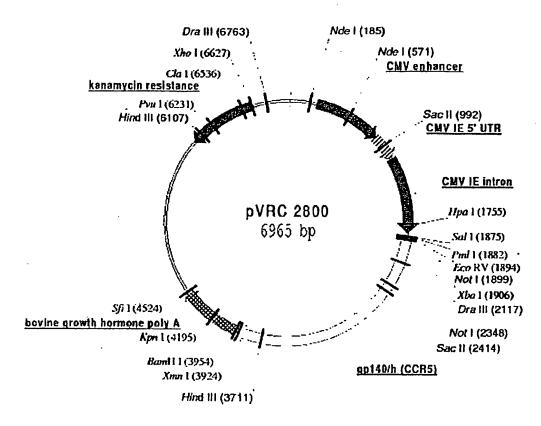


Figure 11

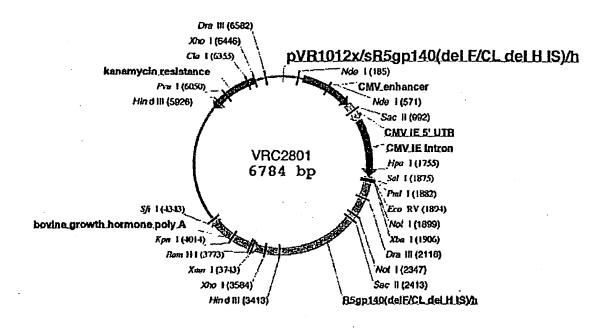


Figure 12

#### pVR1012x/s R5gp145/h

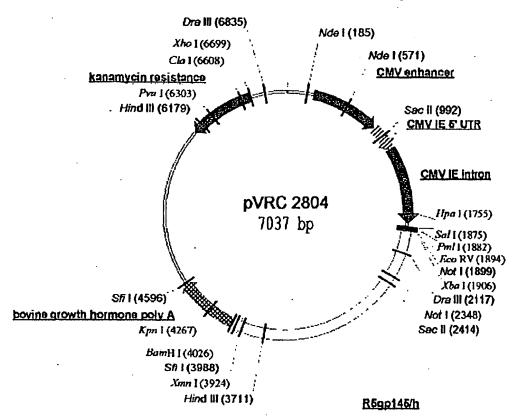


Figure 13

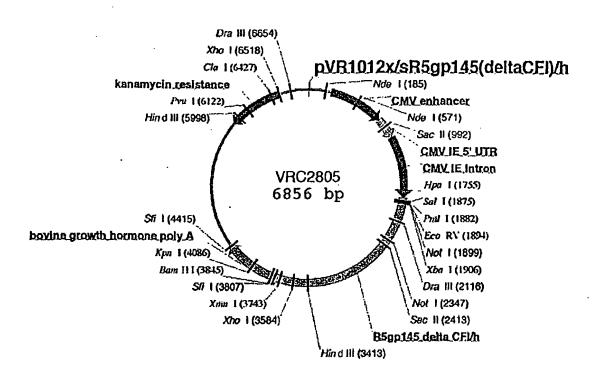


Figure 14

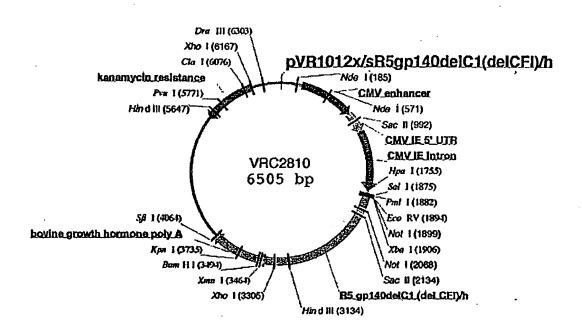


Figure 15

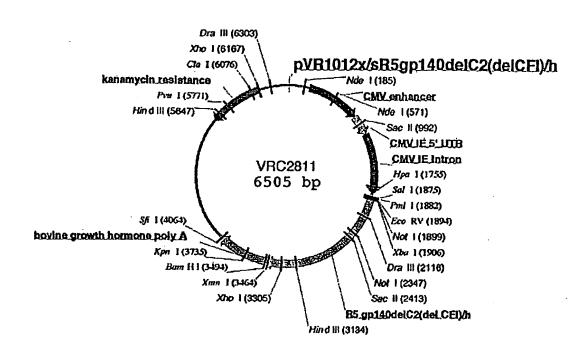


Figure 16

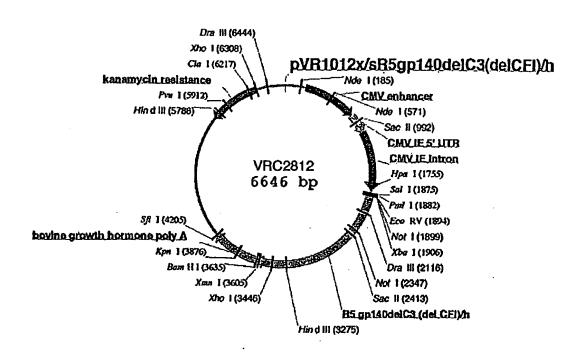


Figure 17

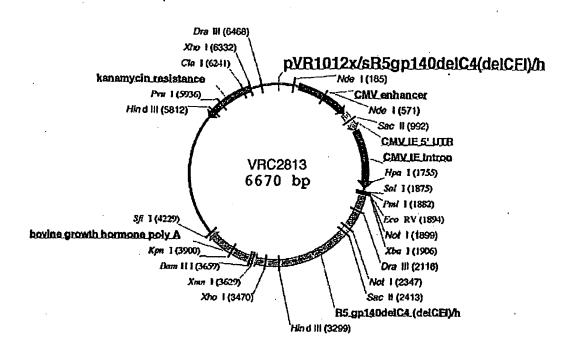


Figure 18

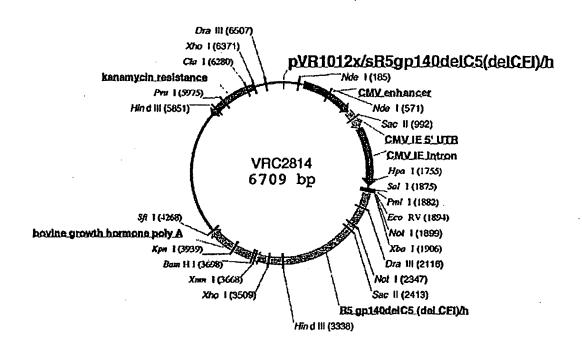


Figure 19

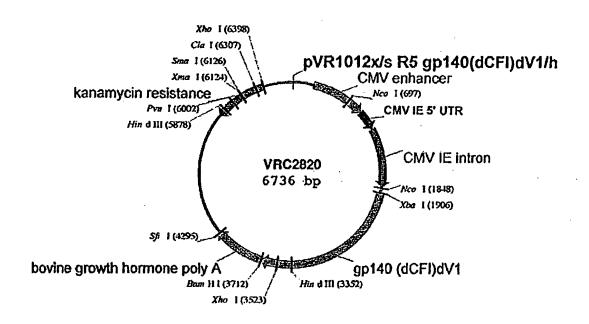


Figure 20

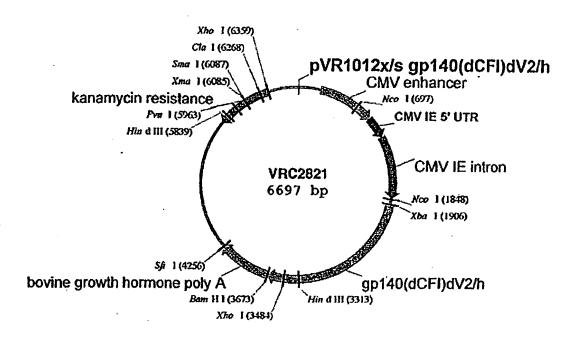


Figure 21

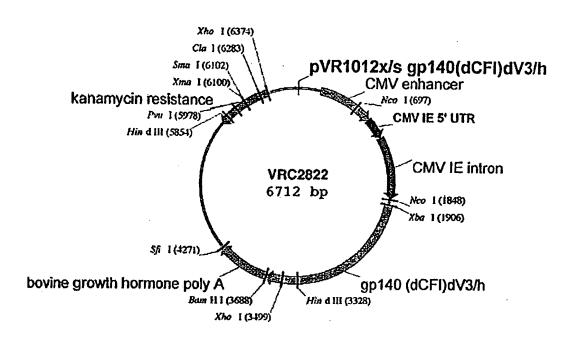


Figure 22

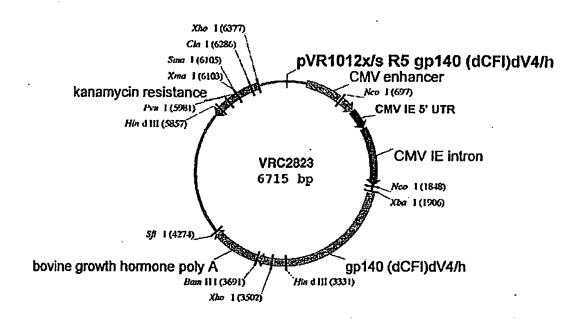


Figure 23

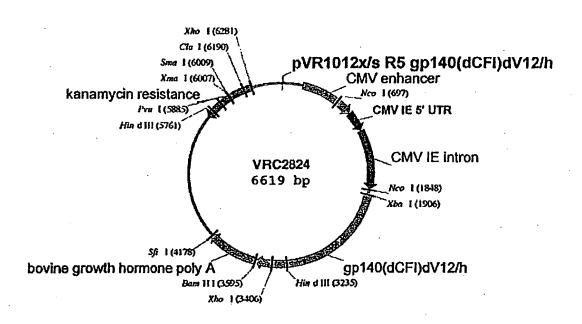


Figure 24

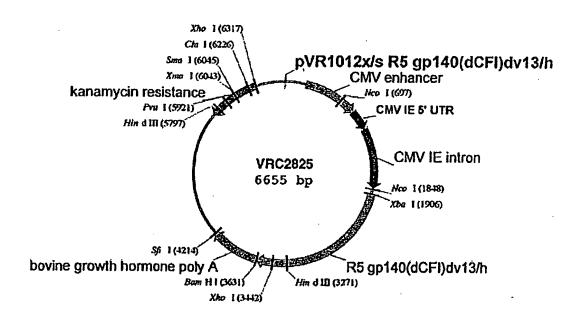


Figure 25

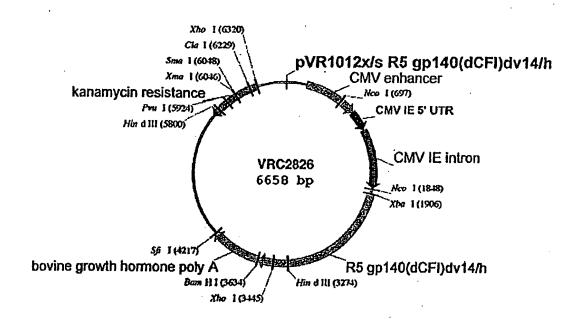


Figure 26

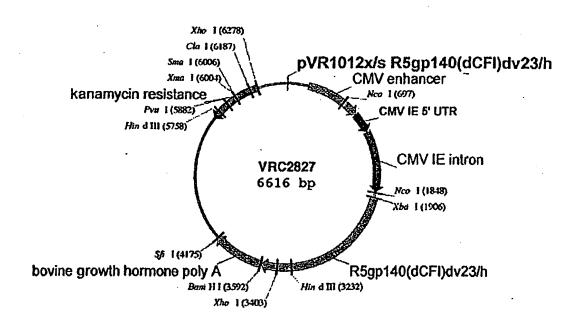


Figure 27

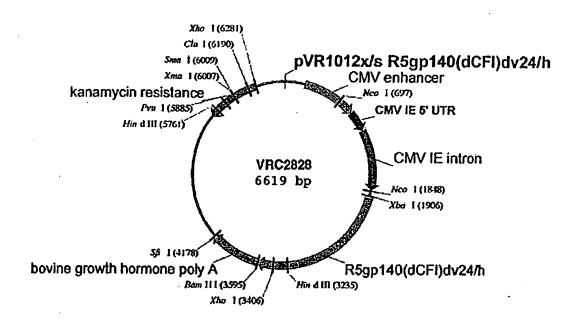


Figure 28

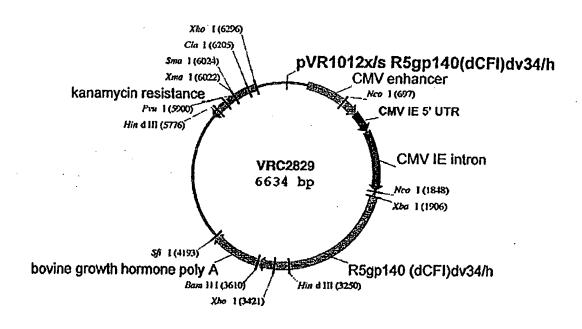


Figure 29

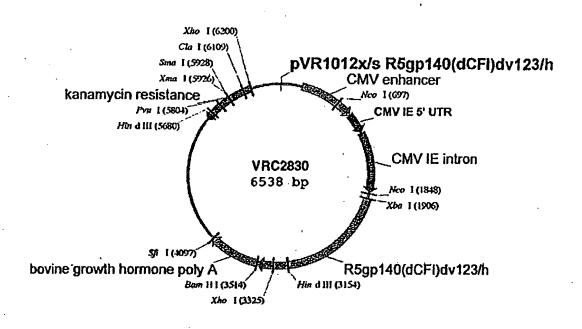


Figure 30

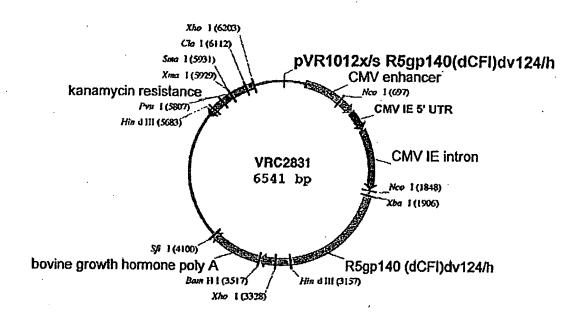


Figure 31

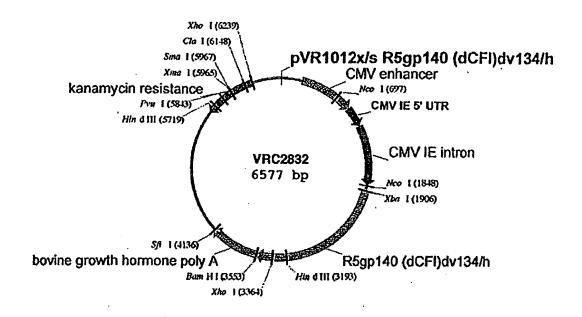


Figure 32

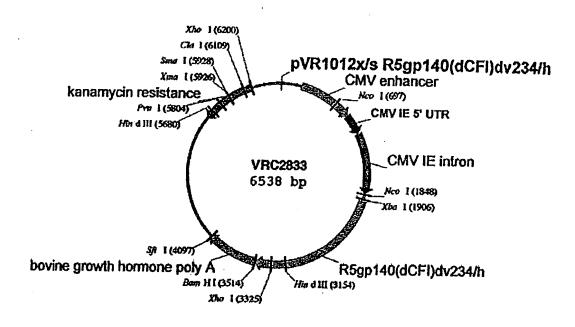


Figure 33

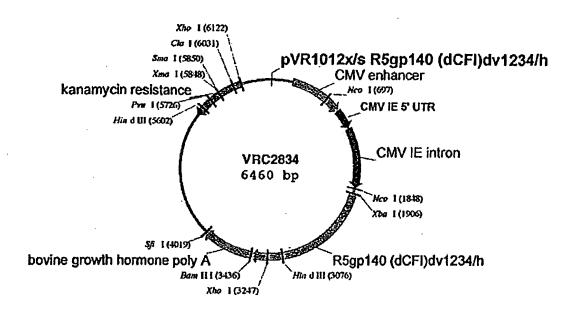


Figure 34

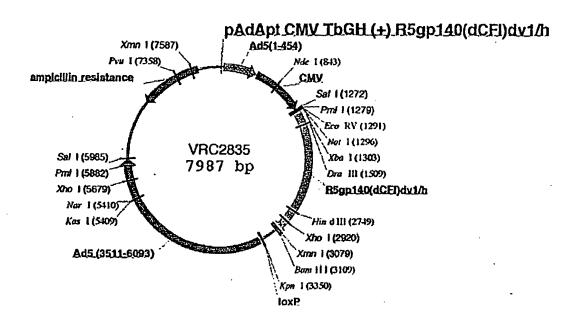


Figure 35

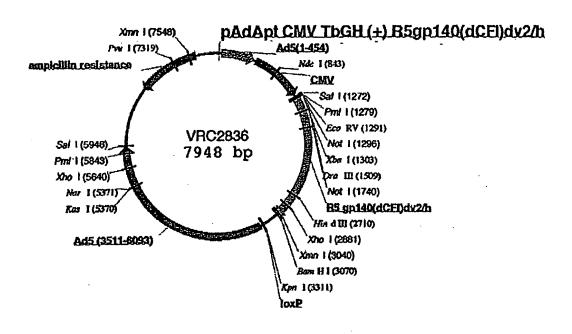


Figure 36

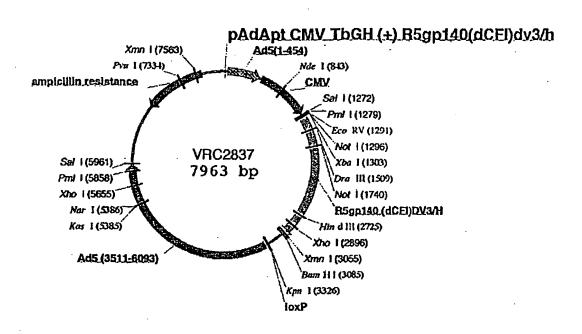


Figure 37

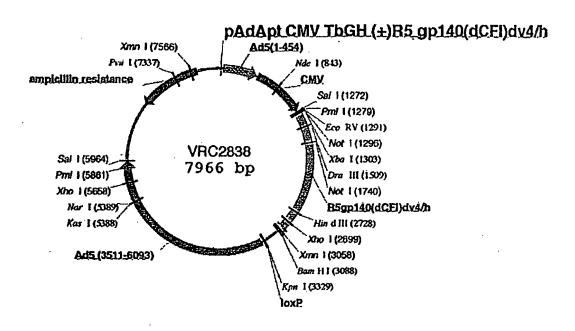


Figure 38

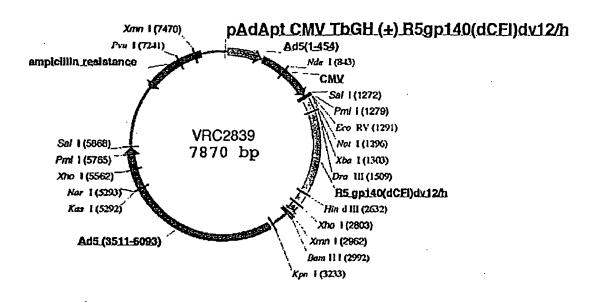


Figure 39

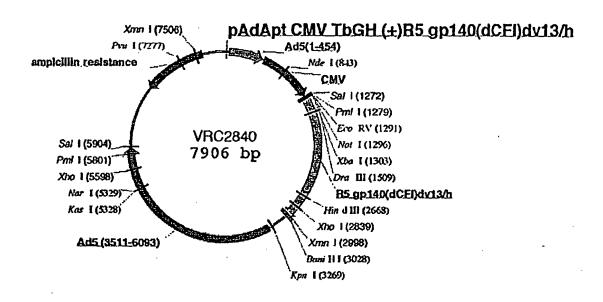


Figure 40

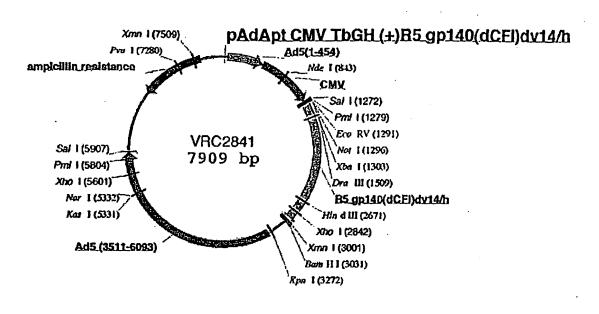


Figure 41

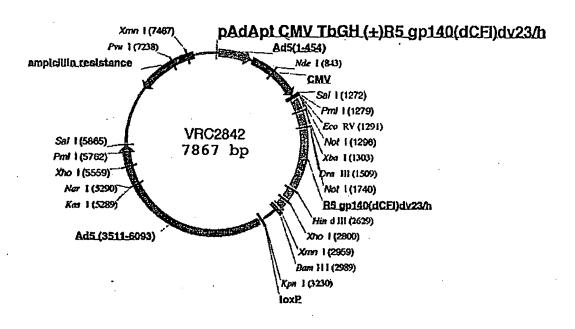


Figure 42

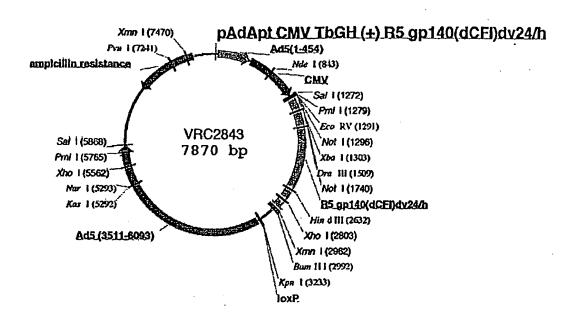


Figure 43

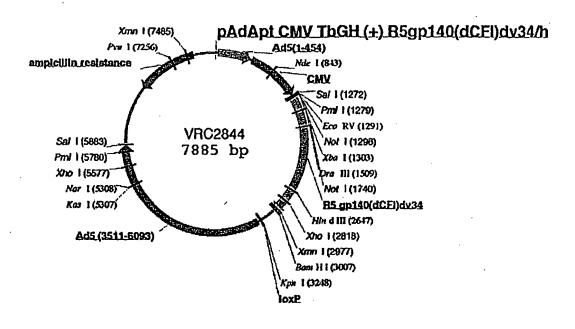


Figure 44

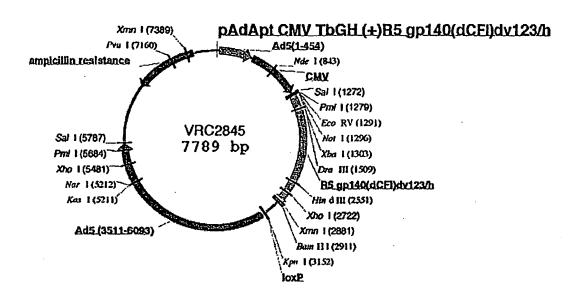


Figure 45

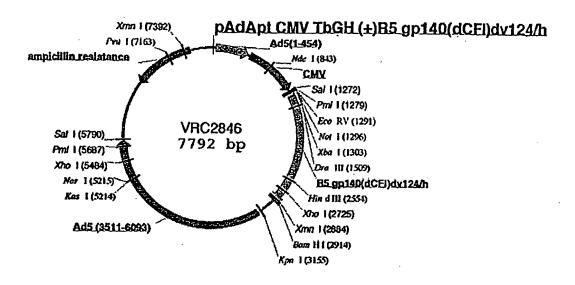


Figure 46

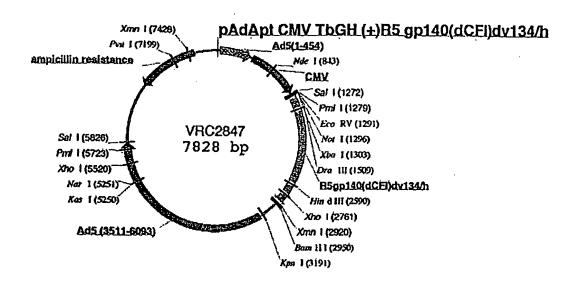


Figure 47

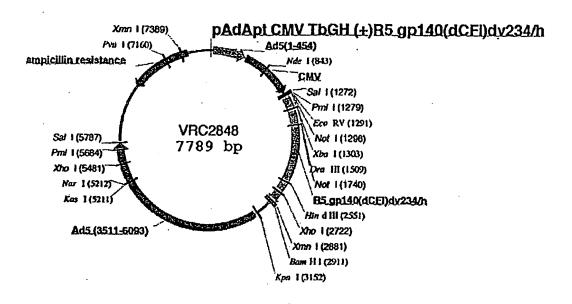


Figure 48

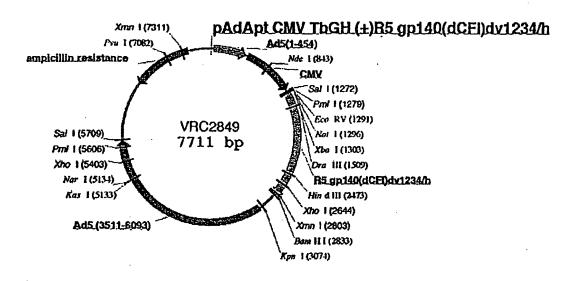


Figure 49

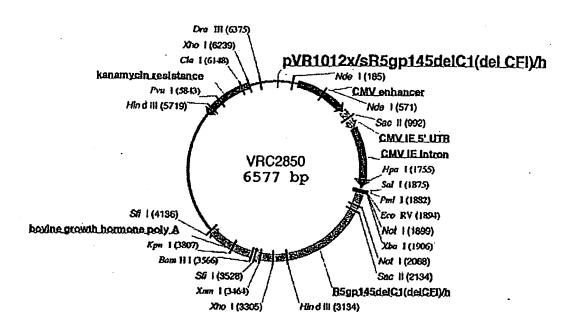


Figure 50

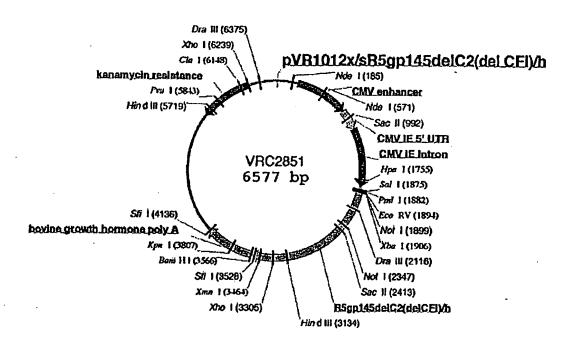


Figure 51

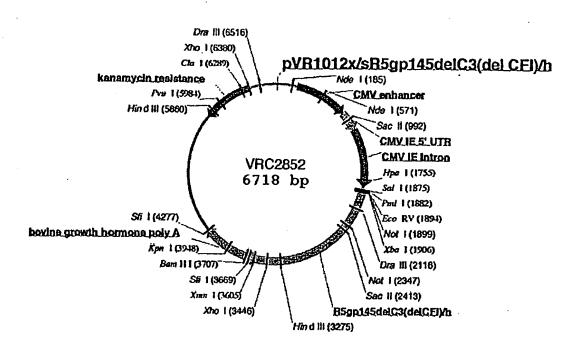


Figure 52

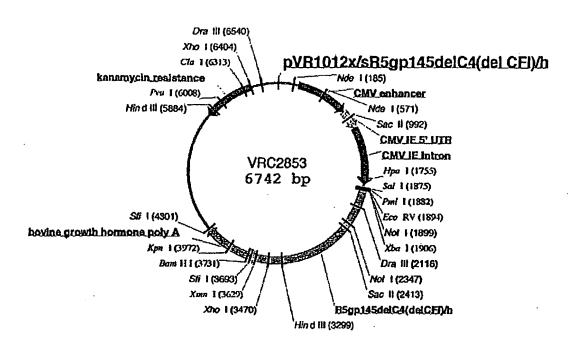


Figure 53

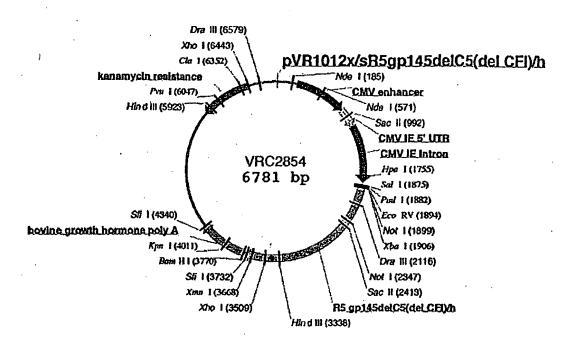


Figure 54

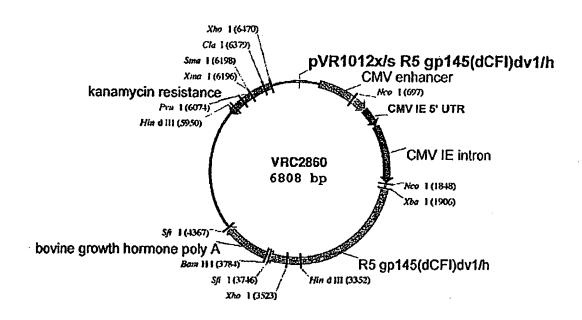


Figure 55

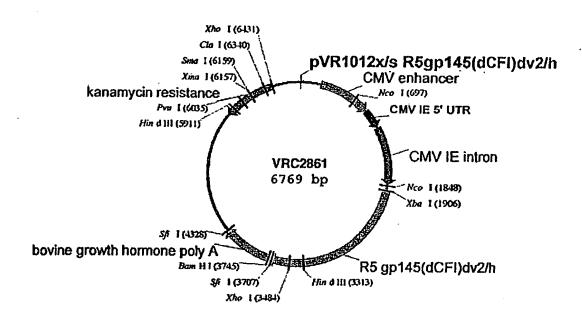


Figure 56

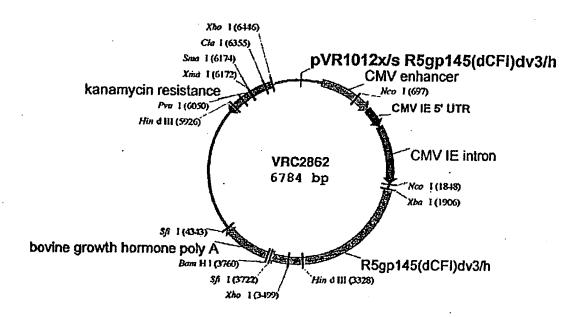


Figure 57

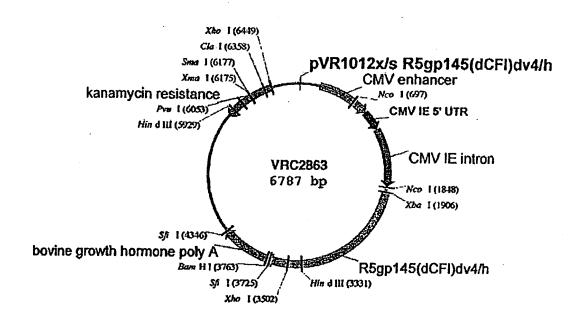


Figure 58

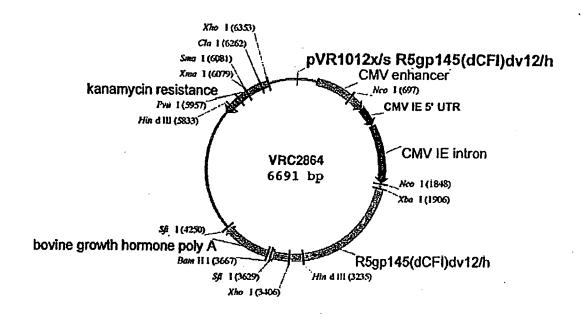


Figure 59

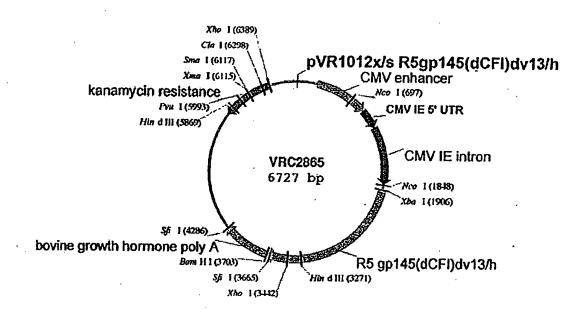


Figure 60

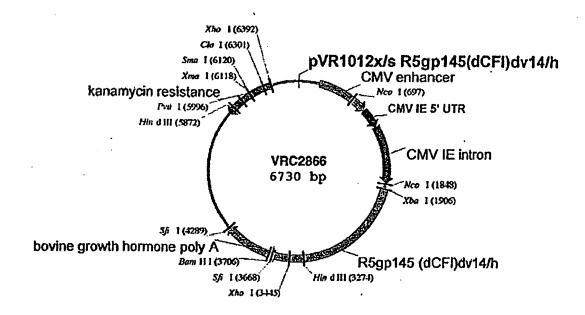


Figure 61

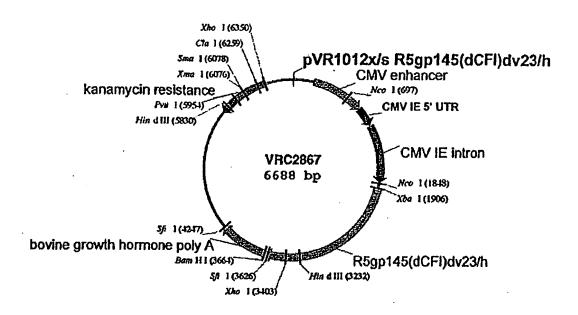


Figure 62

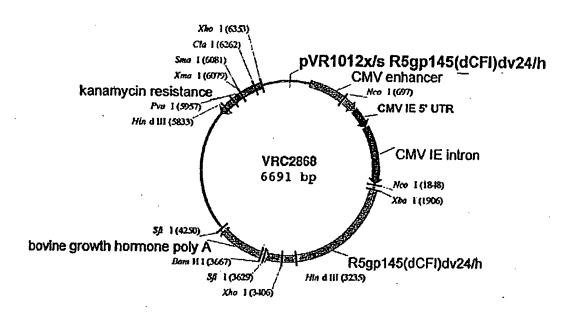


Figure 63

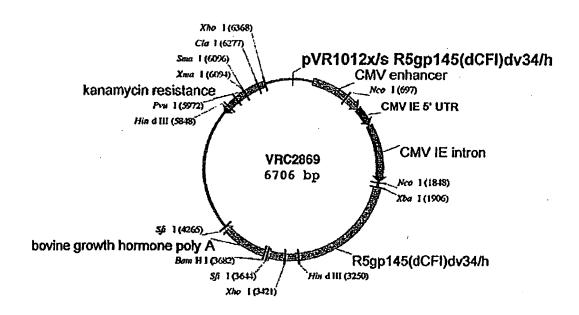


Figure 64

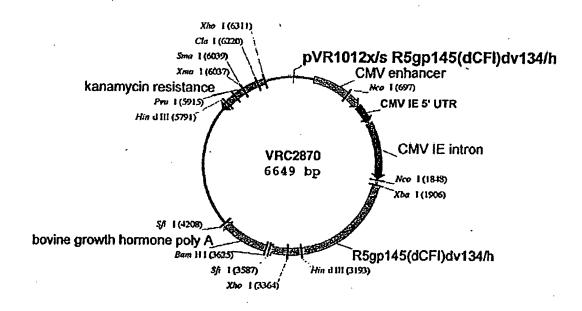


Figure 65

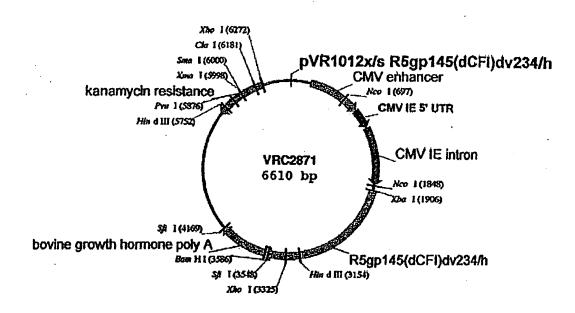


Figure 66

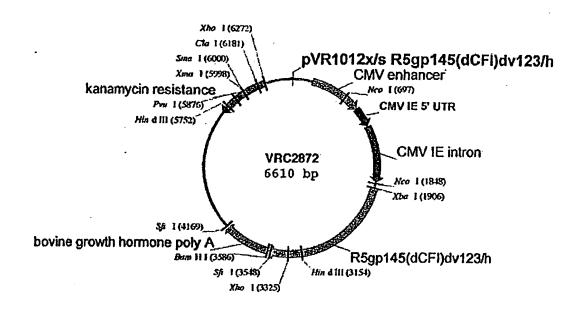
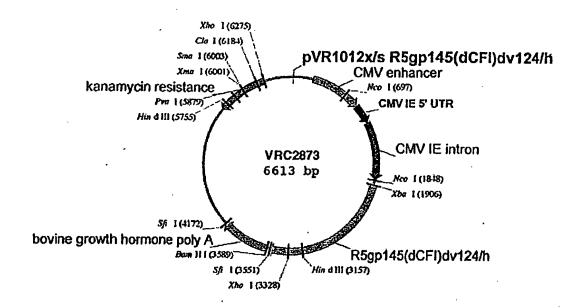


Figure 67



·Figure 68

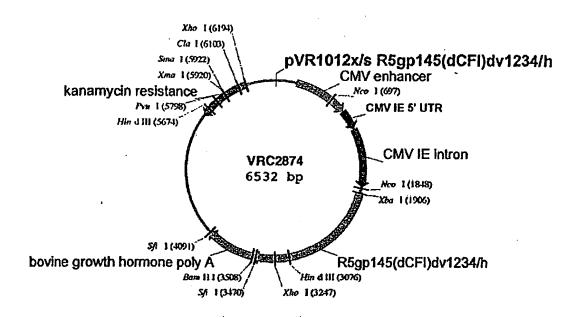


Figure 69

# pVR1012x/s R5gp150/h

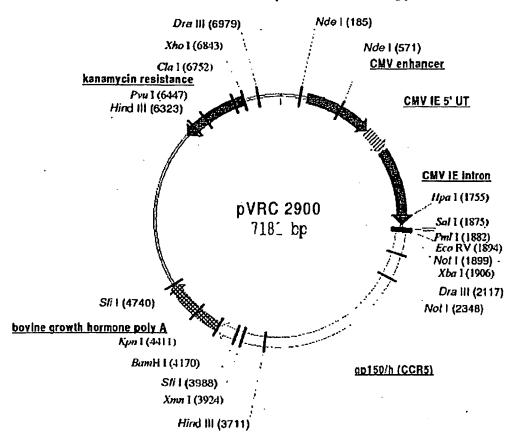


Figure 70

# pVR1012x/sR5gp160/h

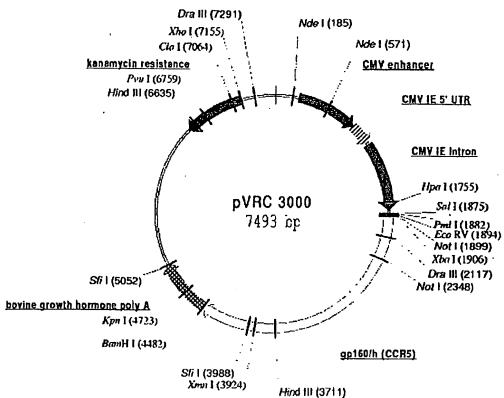


Figure 71

# pVR1012x/s X4gp150/h

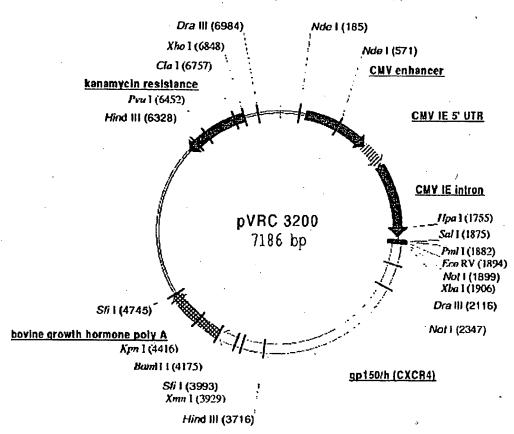


Figure 72

## pVR1012x/s X4gp150(del F/CL del H IS)/h

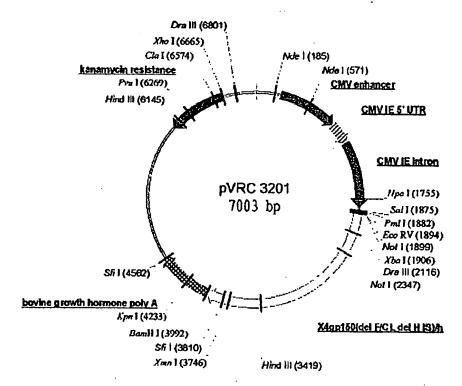


Figure 73

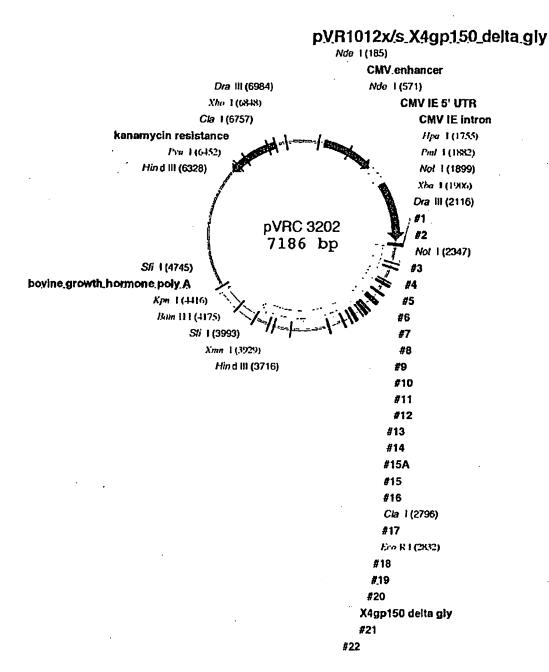


Figure 74

#### pVR1012x/s X4qp150 AB (delta)qly/h Nde I (185) Dra III (6984) Nde 1 (571) Xho I (6848) Cla I (6757) kanamycin resistance CMV enhancer Pru I (6452) Hind III (6328) CMV IE 5' UTR CMY IE Intron **pVRC 3203** Hpa I (1755) Pml I (1882) 7186 bp Eco RV (1894) Not I (1899) Xha I (1906) Ore III (2116) Not1 (2347) Sfi I (4745) bovine growth hormone poly A Cla I (2796) Kpn I (4416) Bamil 1 (4175) Sti I (3993) Xnn I (3929)

Hind III (3716)

Figure 75

X4qp150 AB (delta)qlv/h

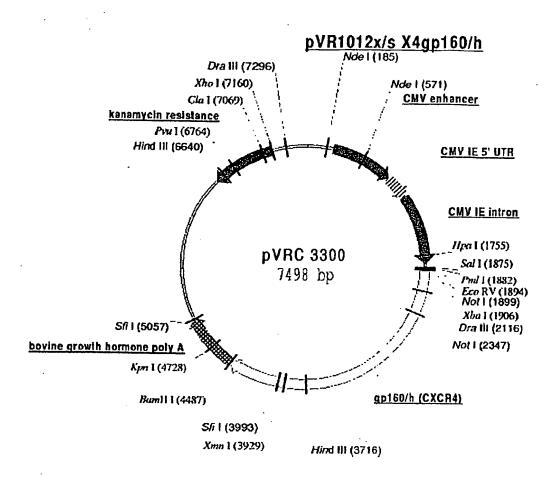


Figure 76

## pVR1012x/s X4ap160(del F/CL del H IS)/h

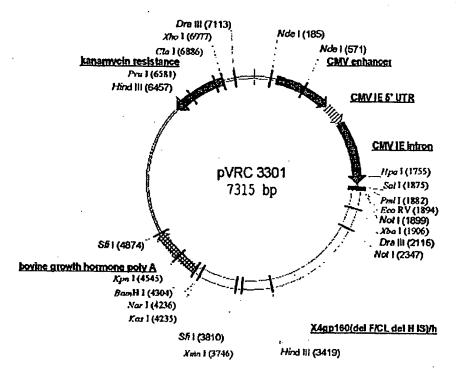


Figure 77

## pVR1012x/s X4gp160 delta gly

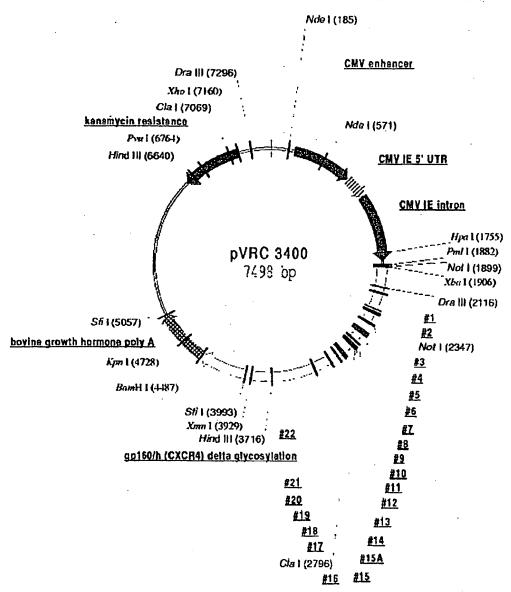


Figure 78

## pVR1012x/s X4qp160 AB Dqly/h

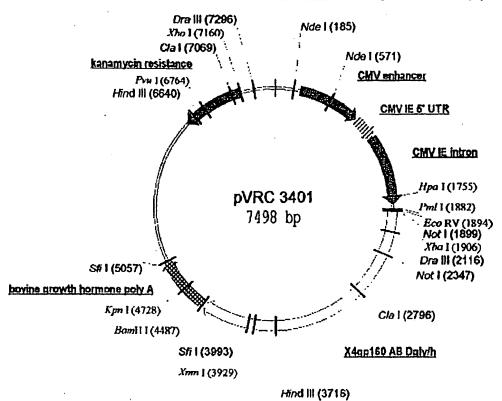


Figure 79

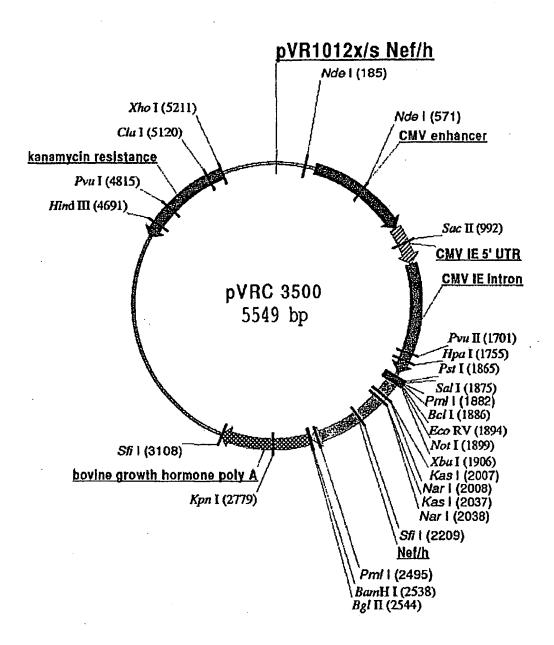


Figure 80

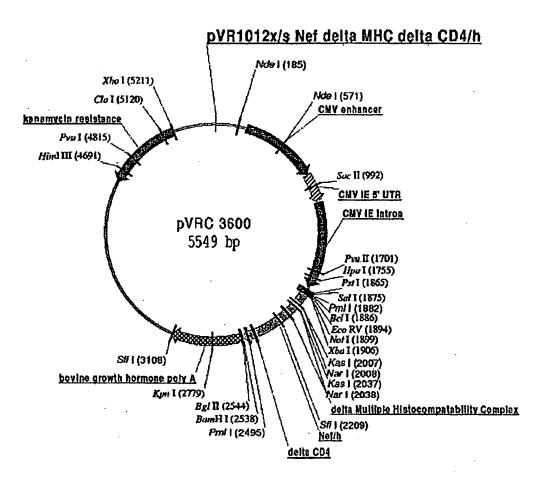


Figure 81

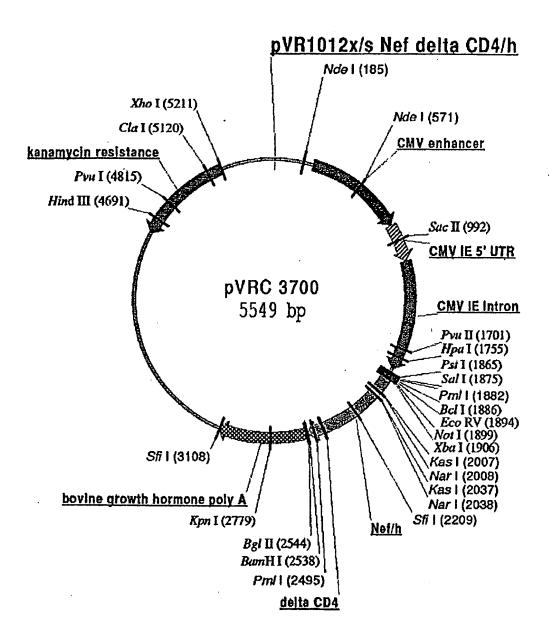


Figure 82

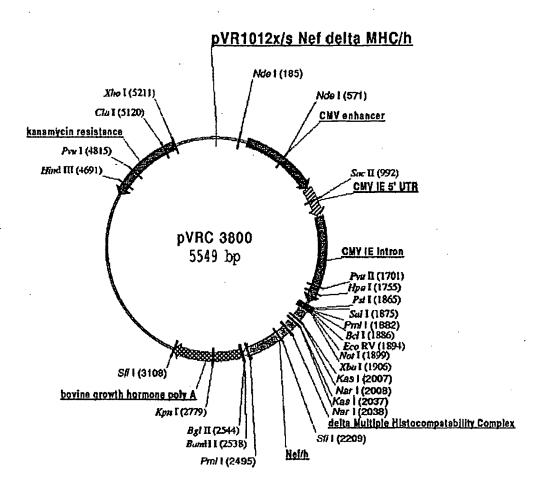


Figure 83

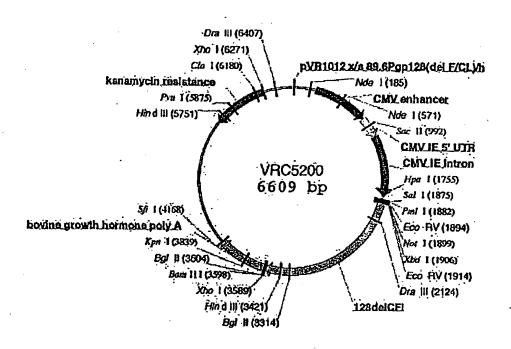


Figure 84

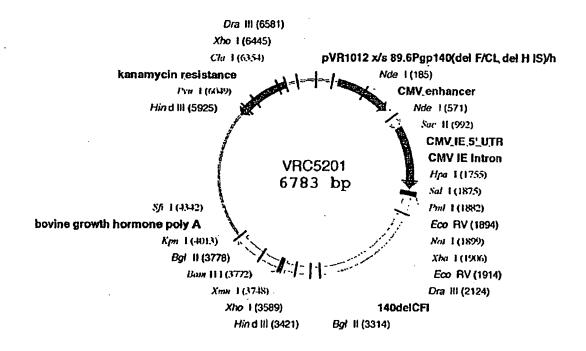


Figure 85

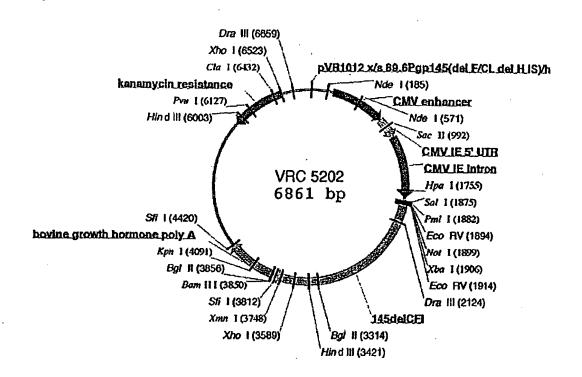


Figure 86

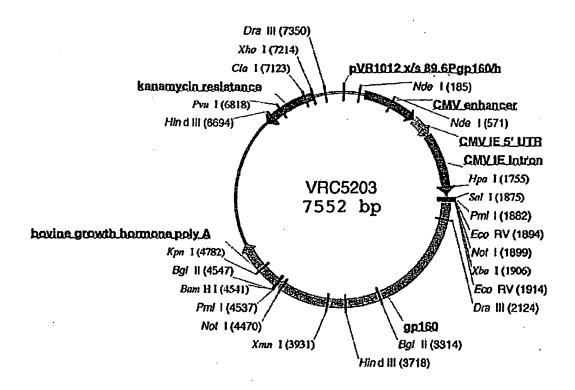


Figure 87

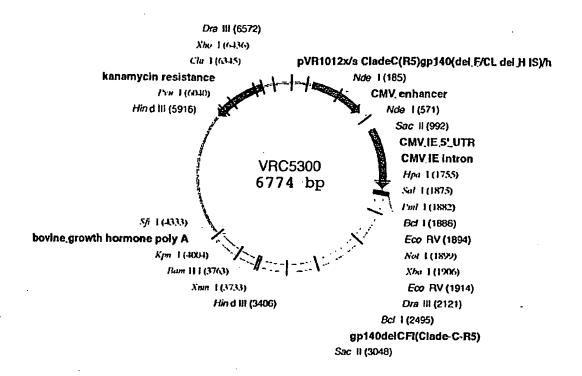


Figure 88

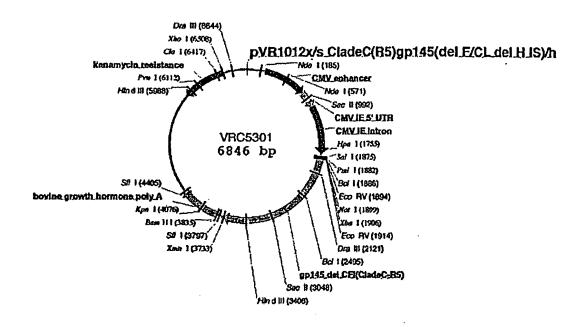


Figure 89

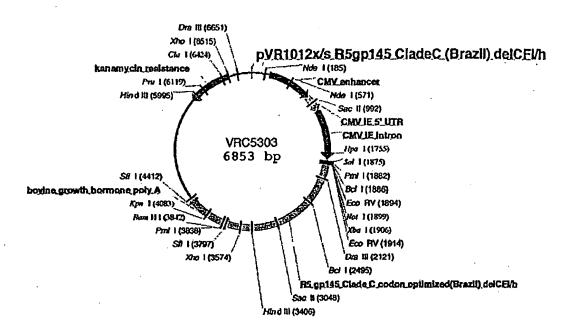


Figure 90

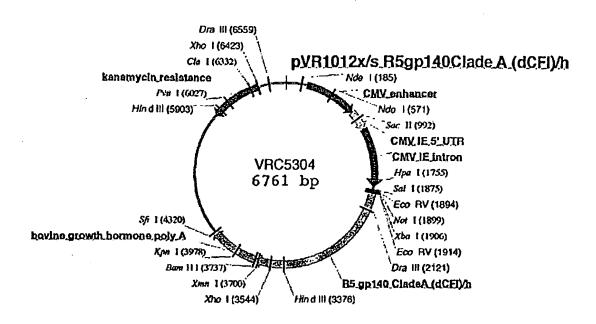


Figure 91

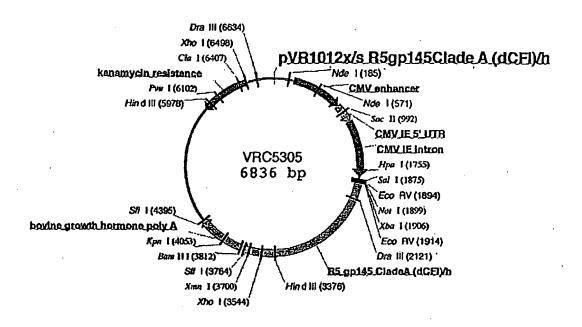


Figure 92

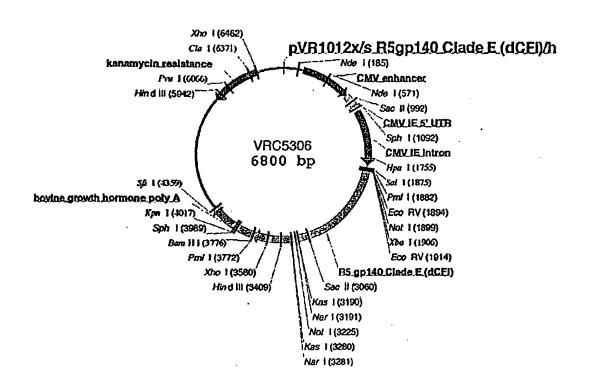


Figure 93

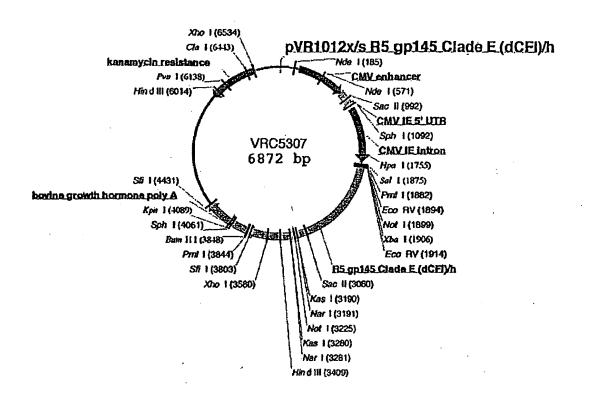


Figure 94

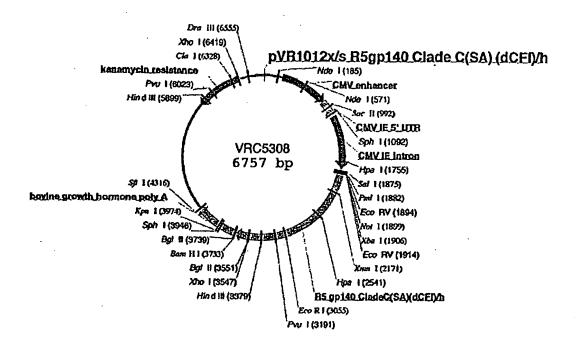


Figure 95

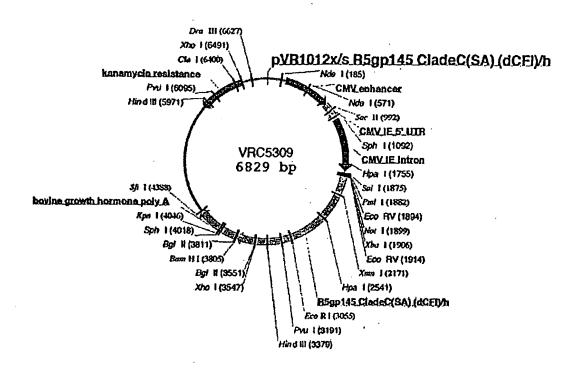


Figure 96

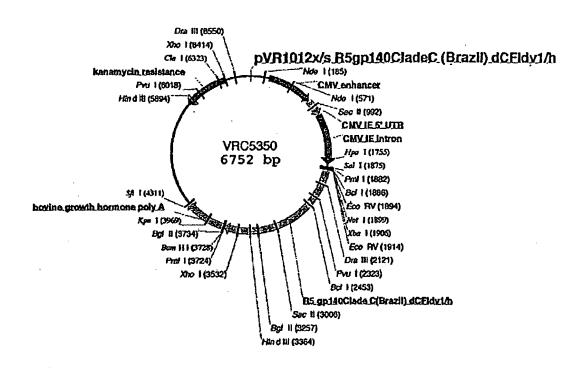


Figure 97

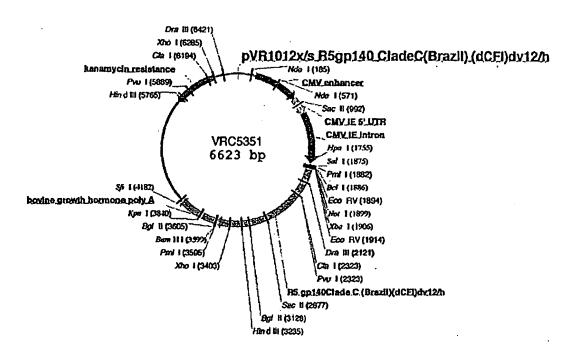


Figure 98

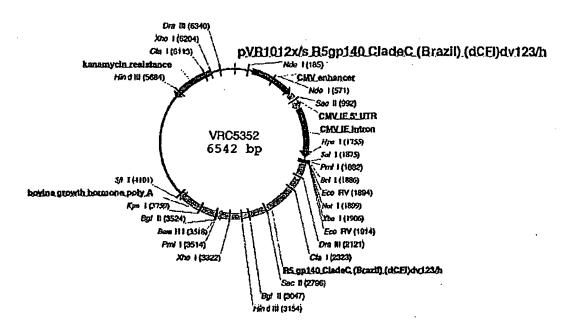


Figure 99

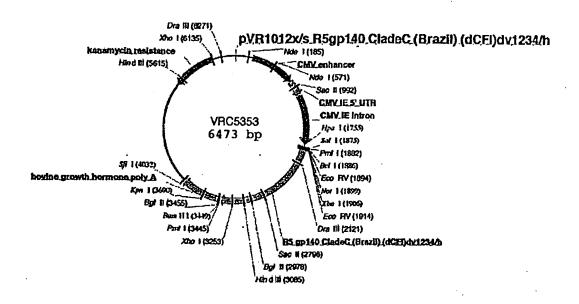


Figure 100

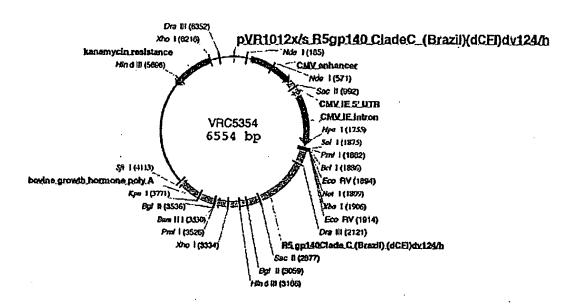


Figure 101

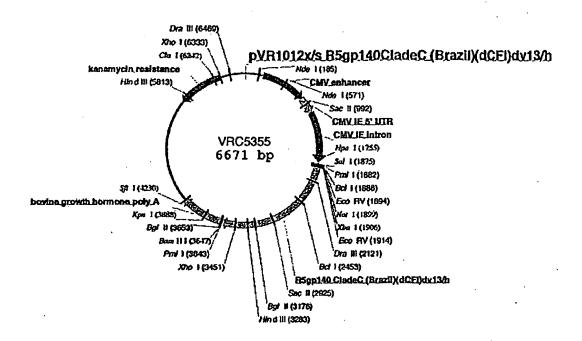


Figure 102

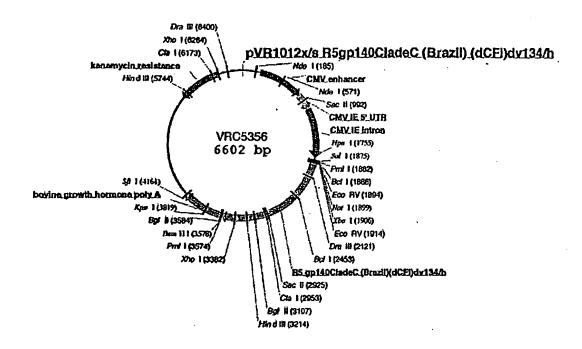


Figure 103

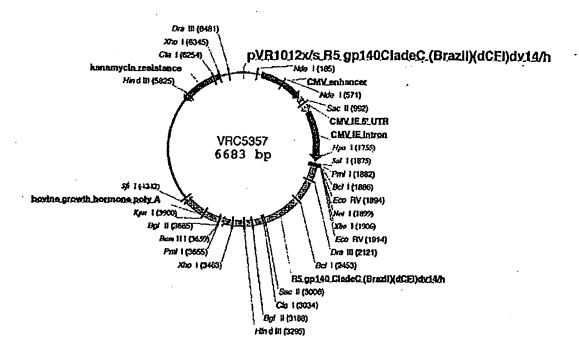


Figure 104

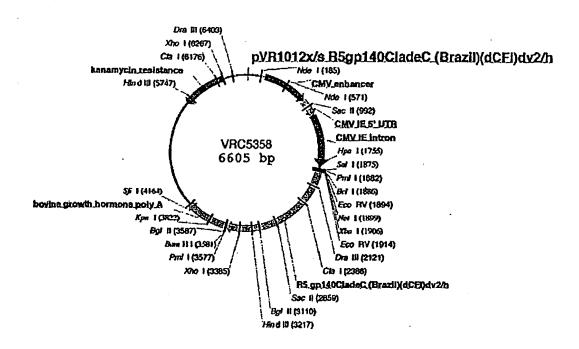


Figure 105

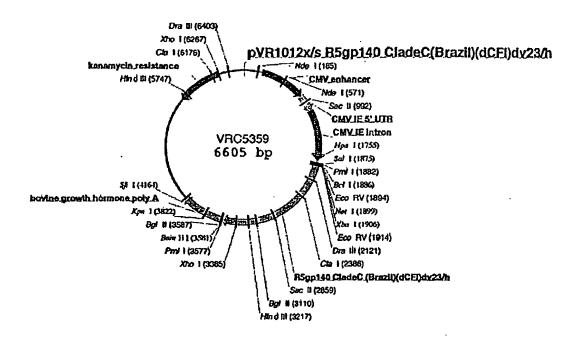


Figure 106

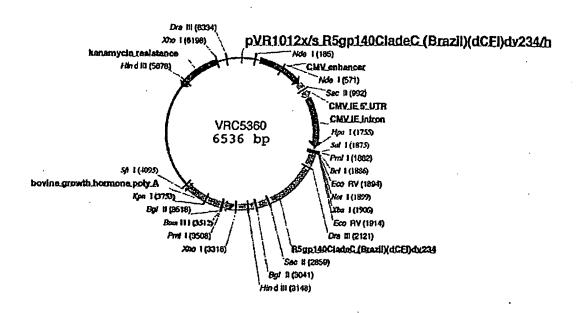


Figure 107

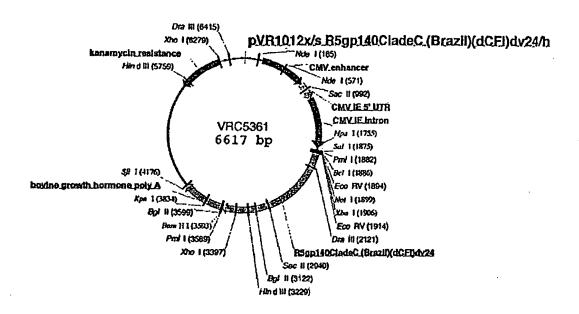


Figure 108

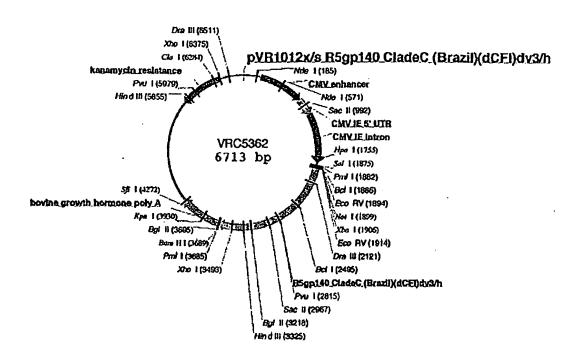


Figure 109

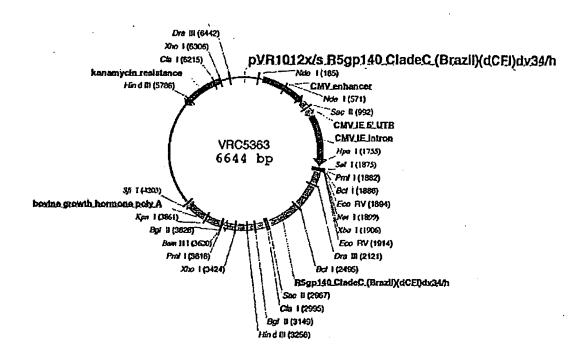


Figure 110

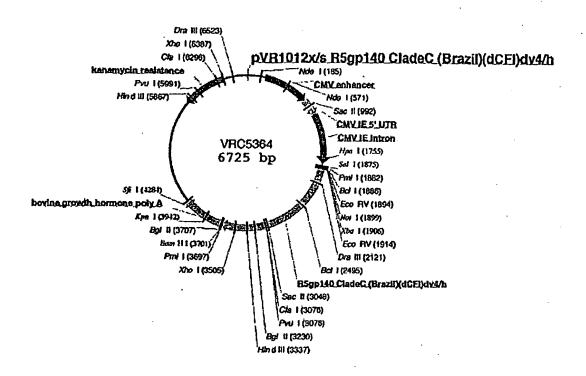


Figure 111

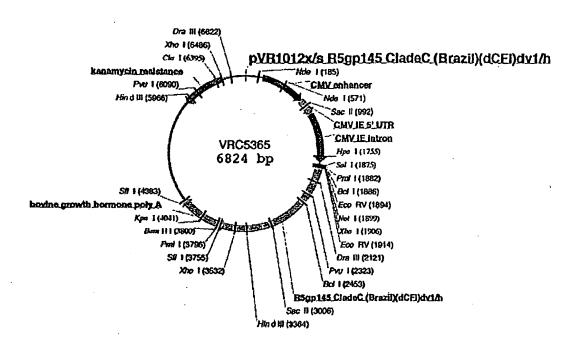


Figure 112

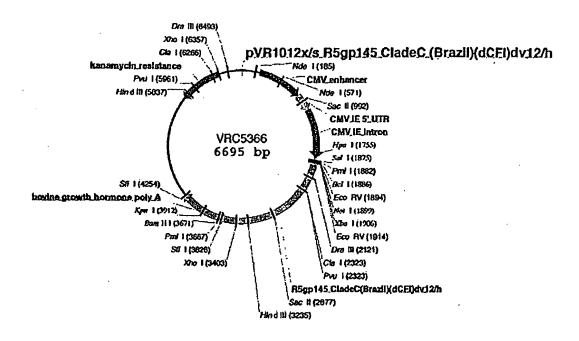


Figure 113

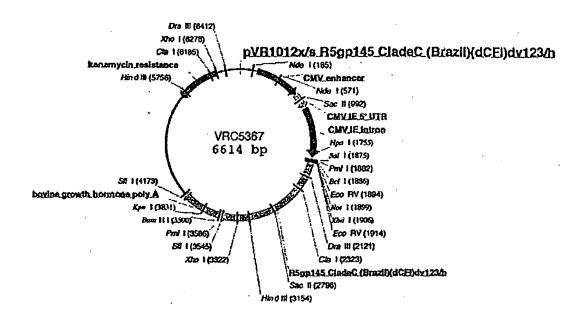


Figure 114

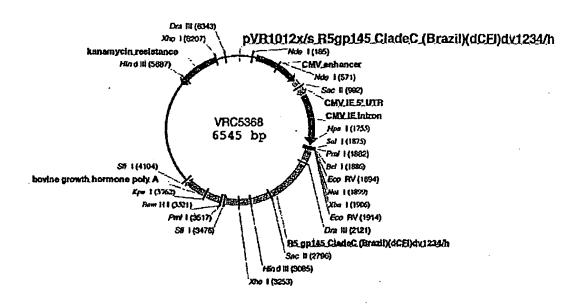


Figure 115

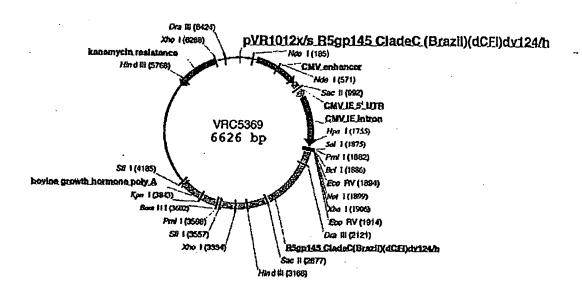


Figure 116

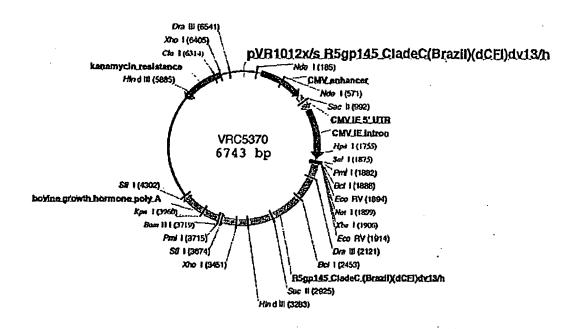


Figure 117

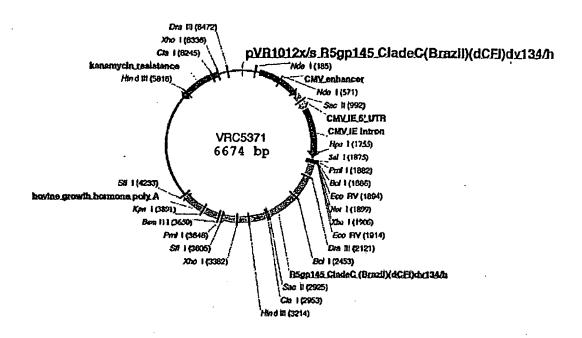


Figure 118

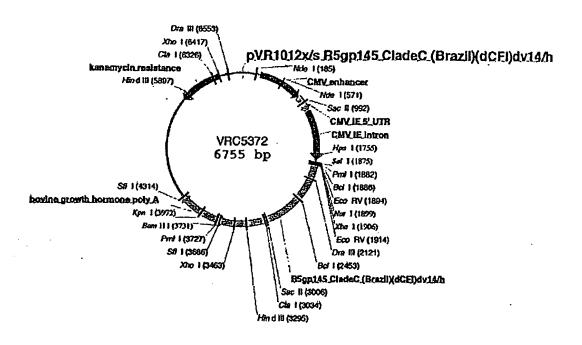


Figure 119

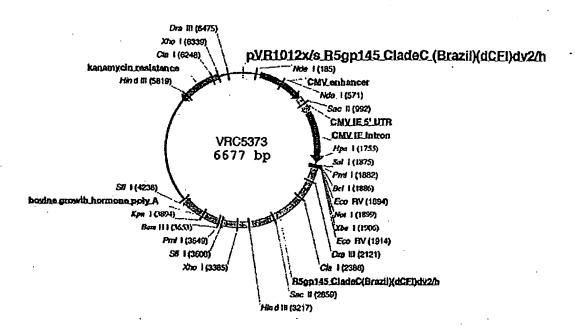


Figure 120

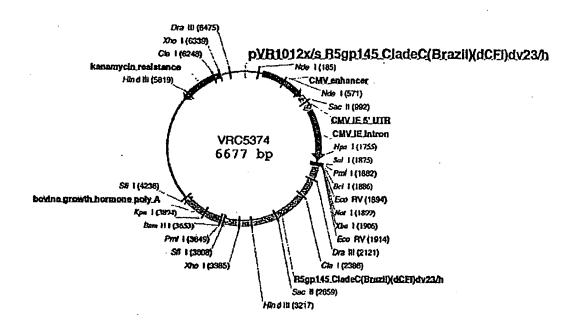


Figure 121

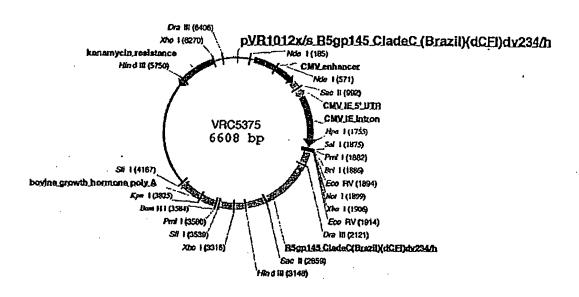


Figure 122

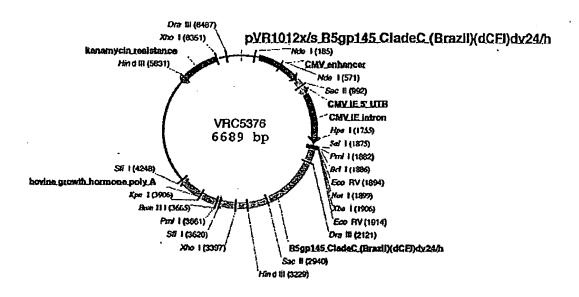


Figure 123

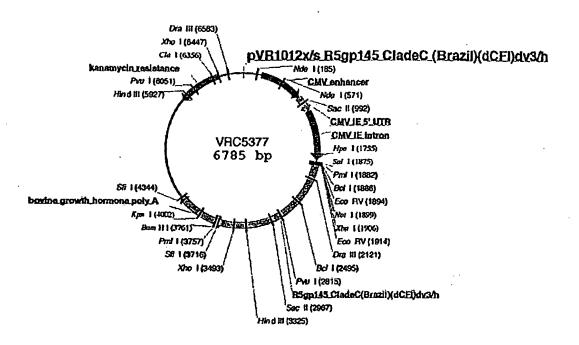


Figure 124

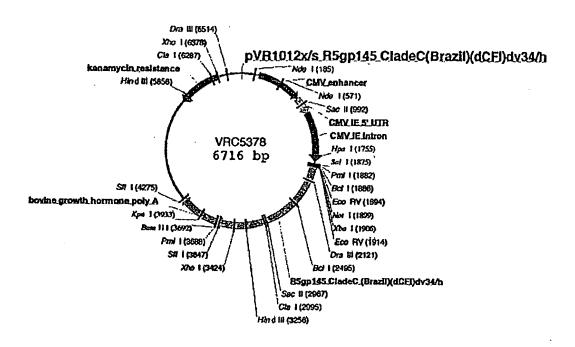


Figure 125

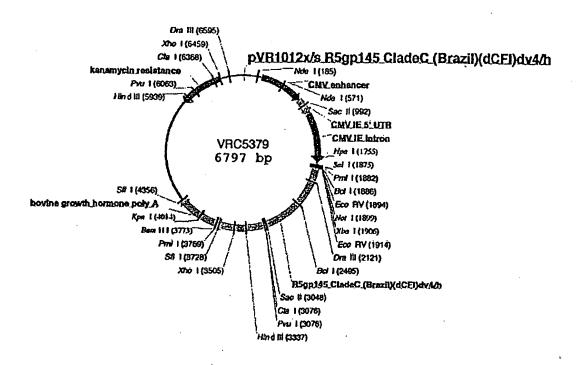


Figure 126

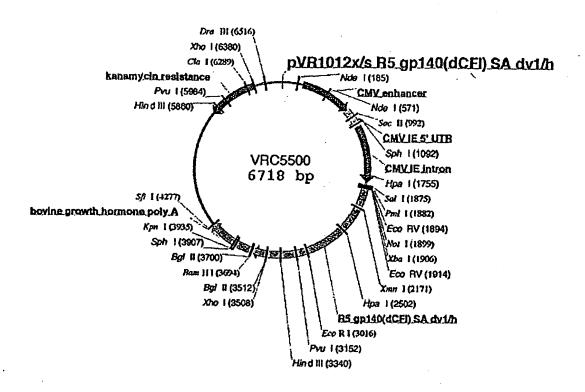


Figure 127

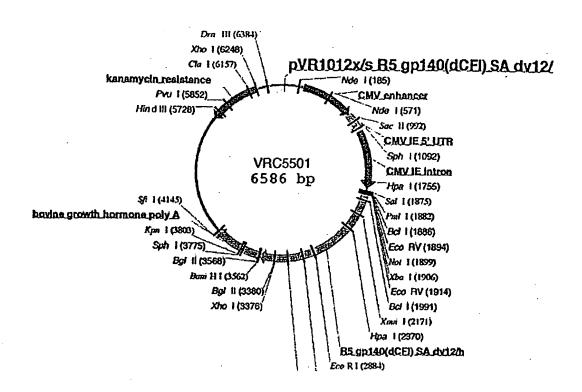


Figure 128

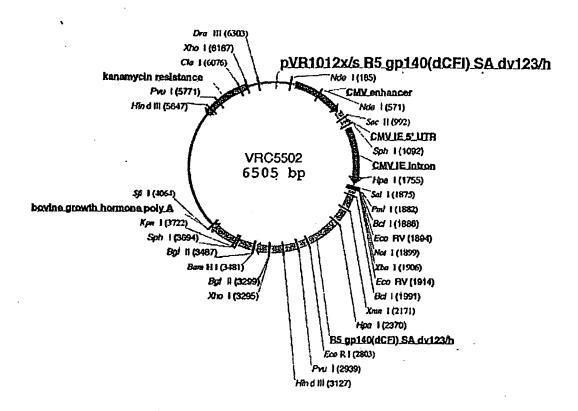


Figure 129

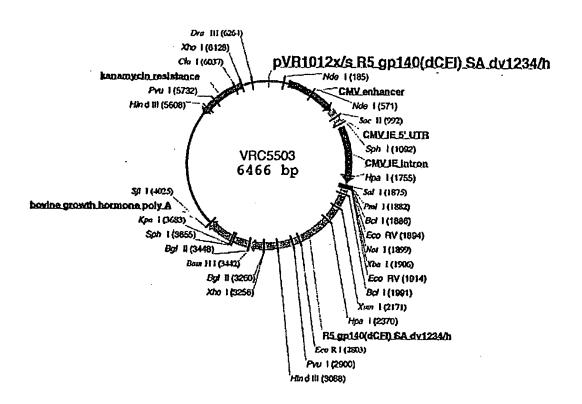


Figure 130

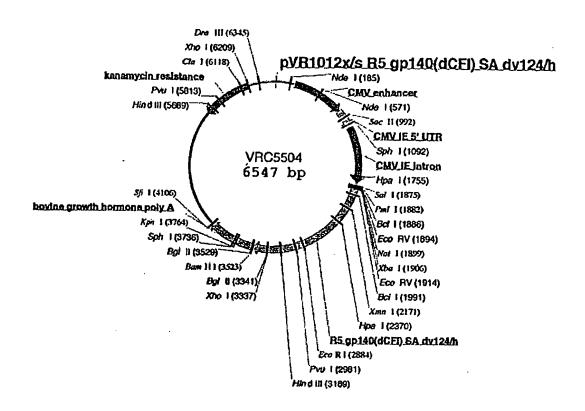


Figure 131

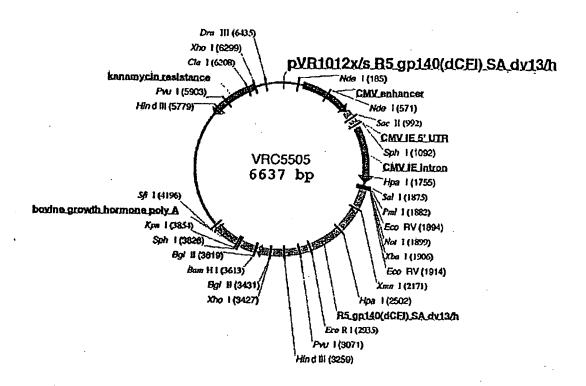


Figure 132

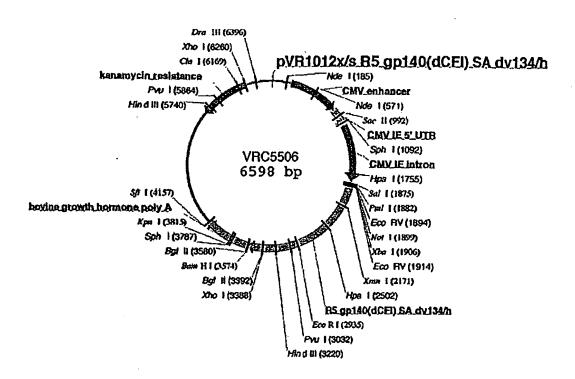


Figure 133

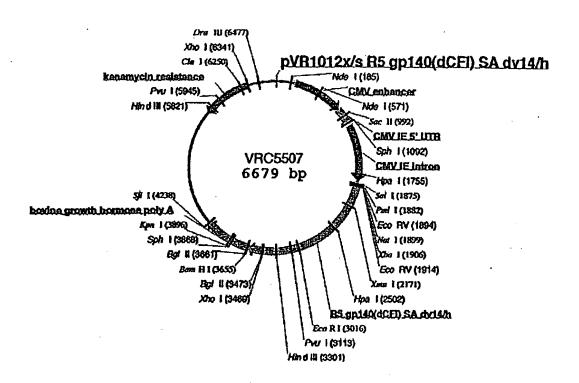


Figure 134

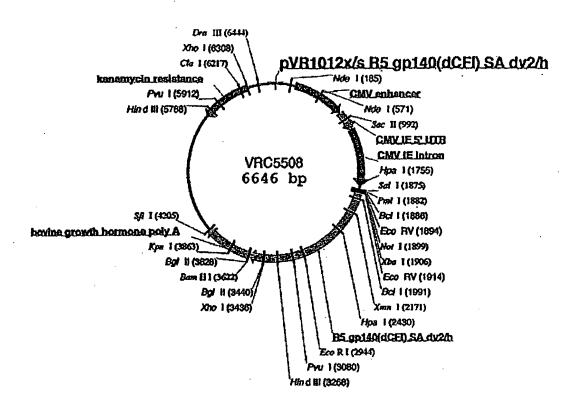


Figure 135

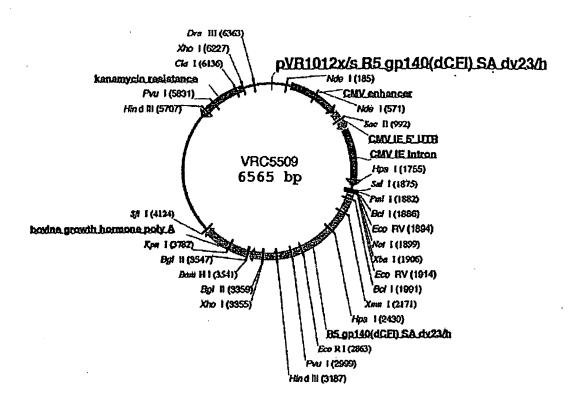


Figure 136

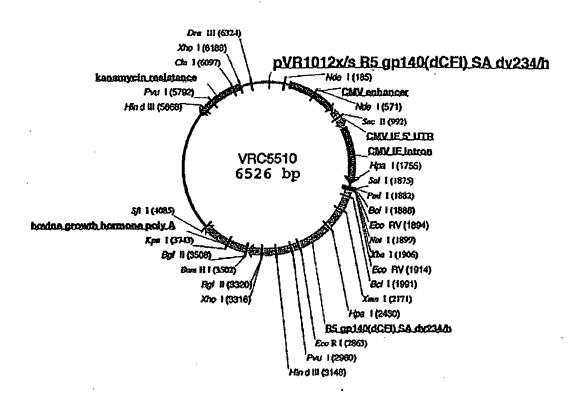


Figure 137

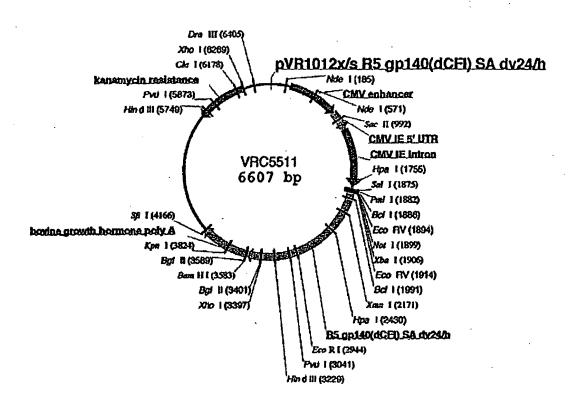


Figure 138

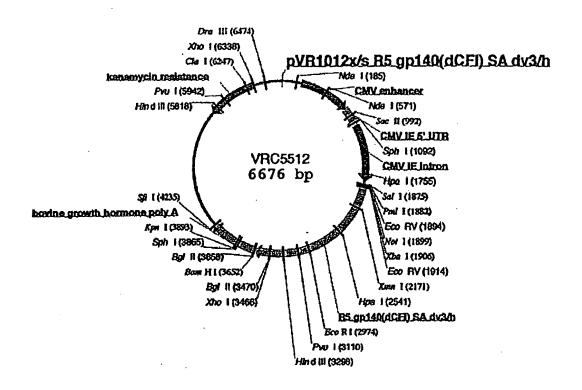


Figure 139

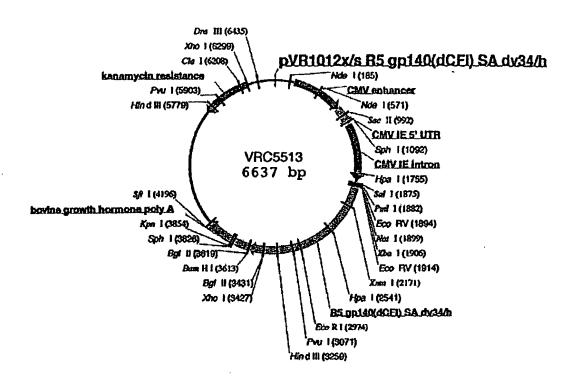


Figure 140

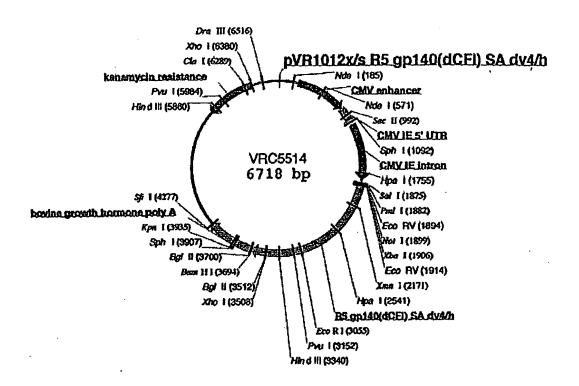


Figure 141

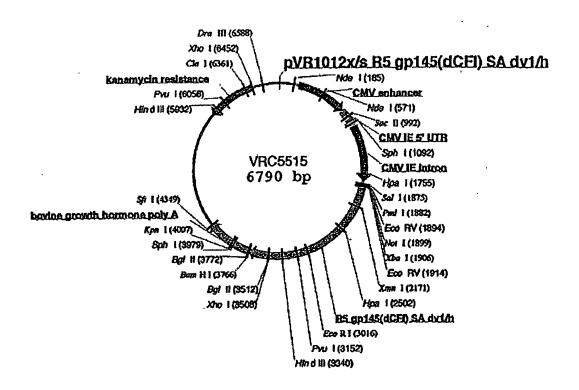


Figure 142

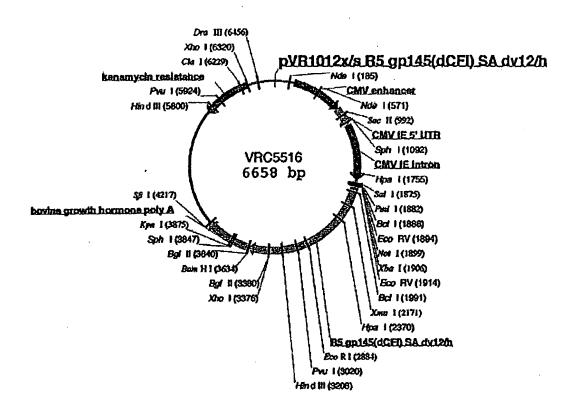


Figure 143

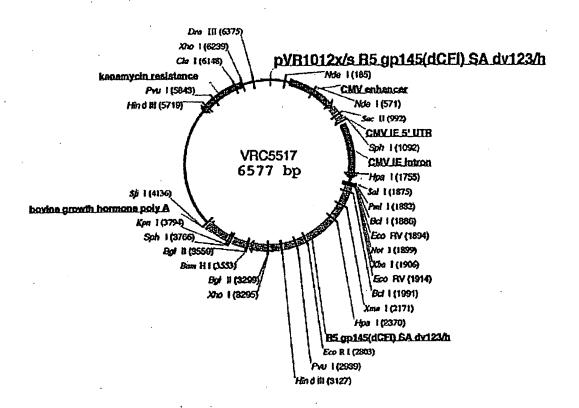


Figure 144

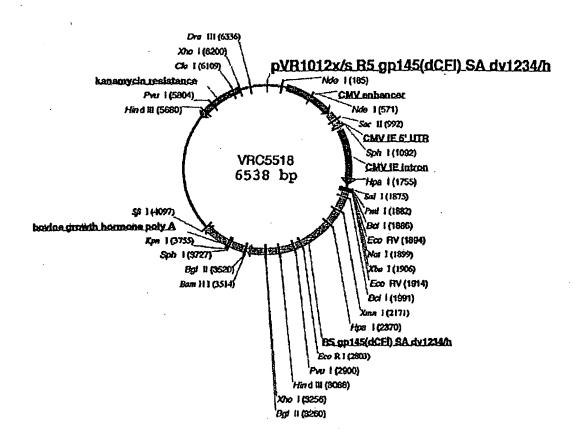


Figure 145

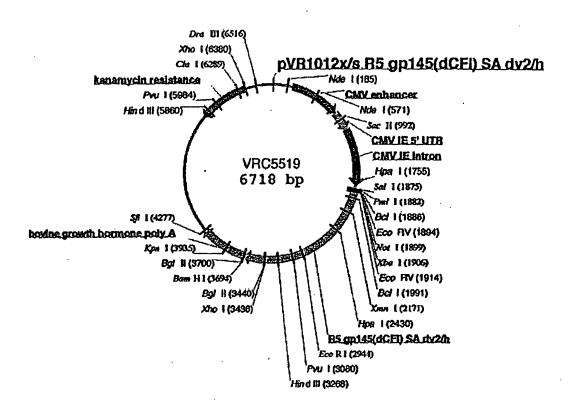


Figure 146

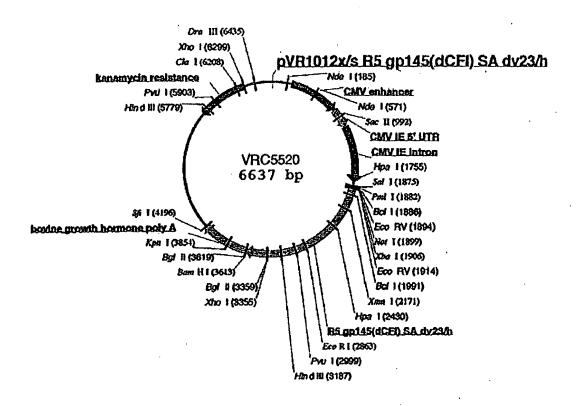
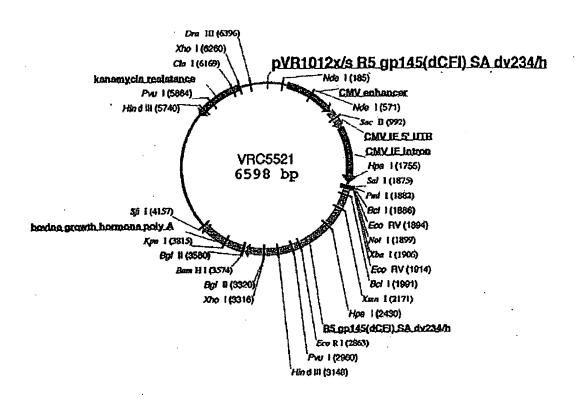


Figure 147



· Figure 148

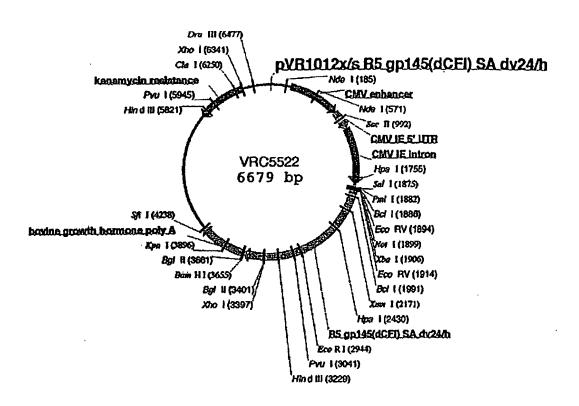


Figure 149

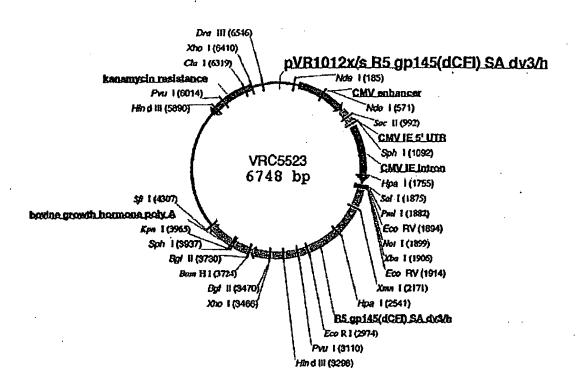


Figure 150

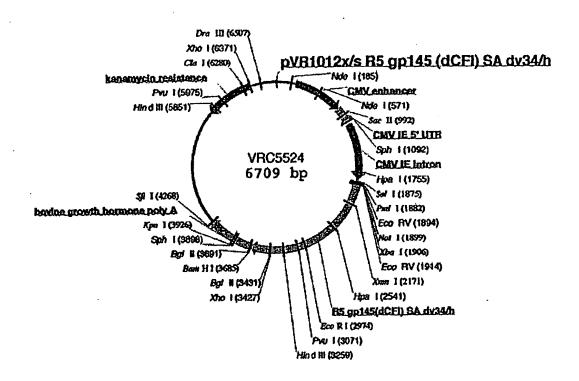


Figure 151

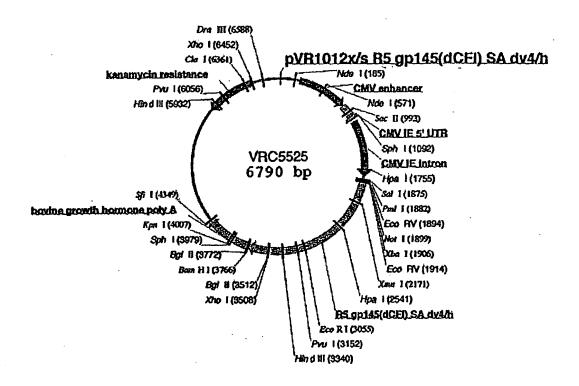


Figure 152

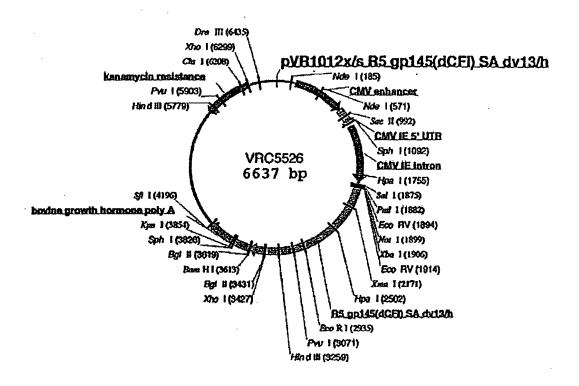


Figure 153

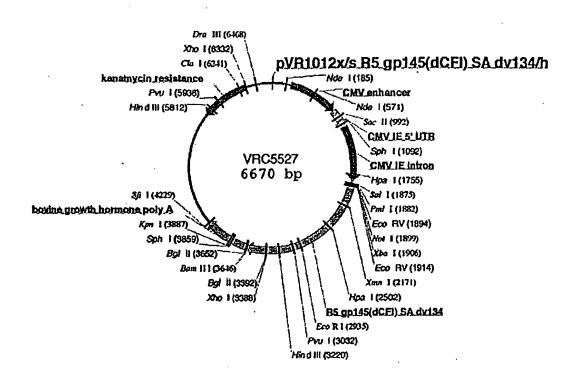


Figure 154

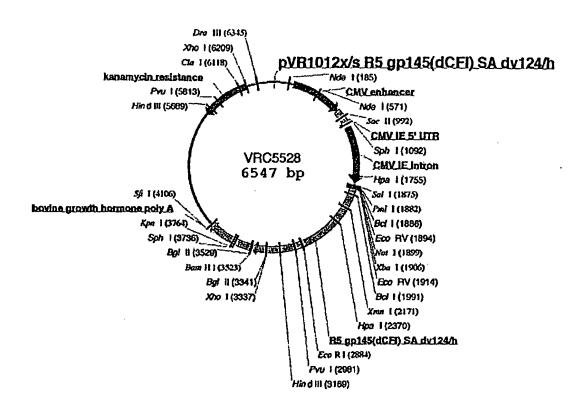


Figure 155

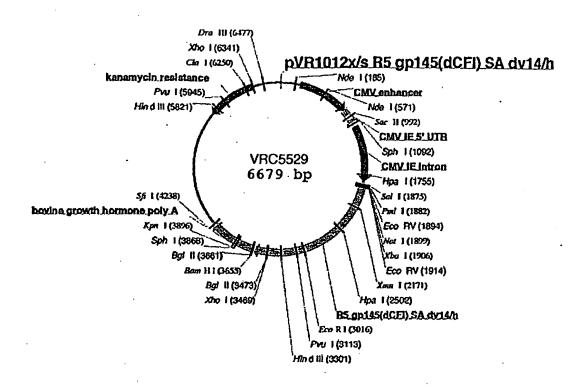


Figure 156

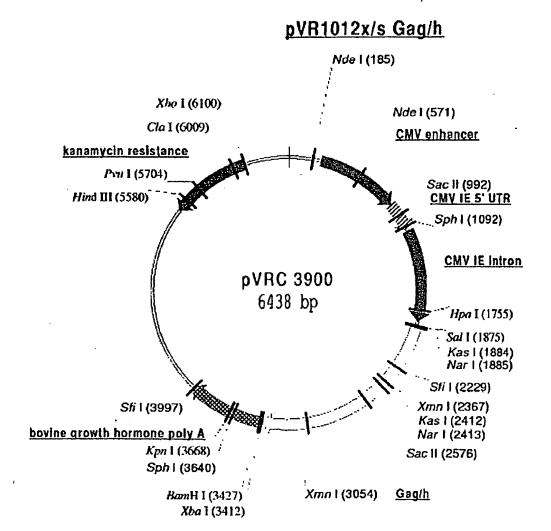


Figure 157

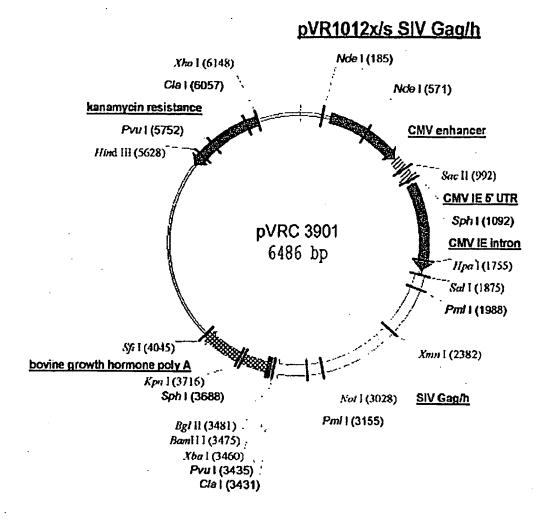


Figure 158

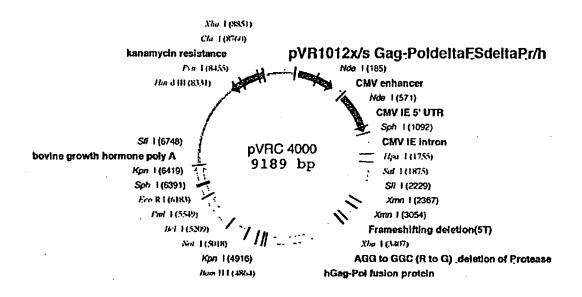


Figure 159

## pVR1012x/s SIV Gag-Pol/h

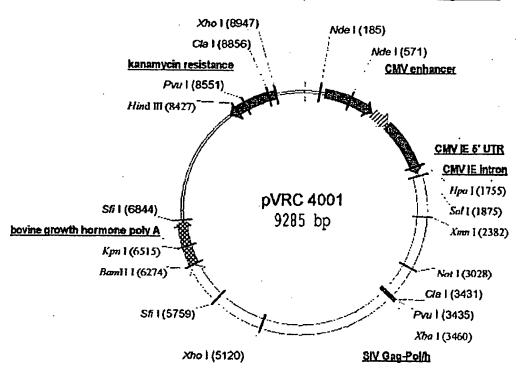


Figure 160

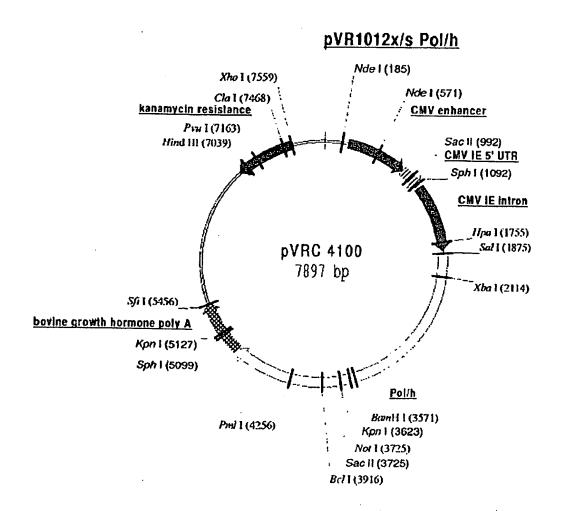


Figure 161

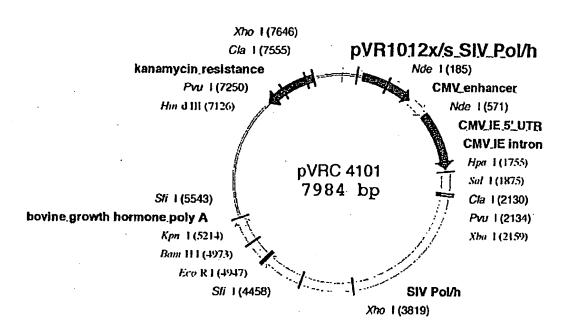


Figure 162

# pVR1012x/s Gag(fs)Pol/h

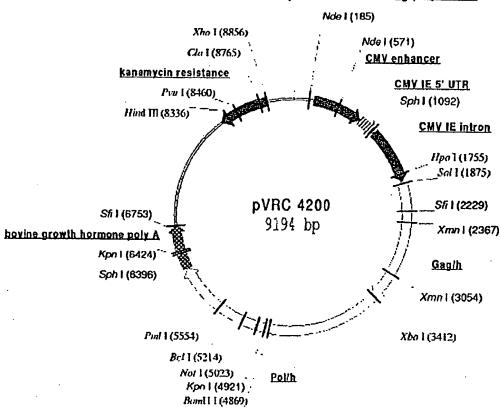


Figure 163

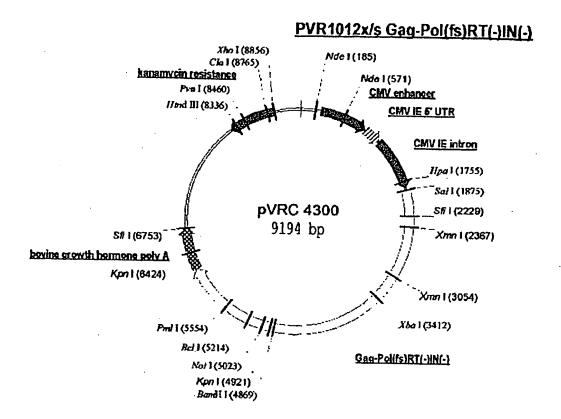


Figure 164

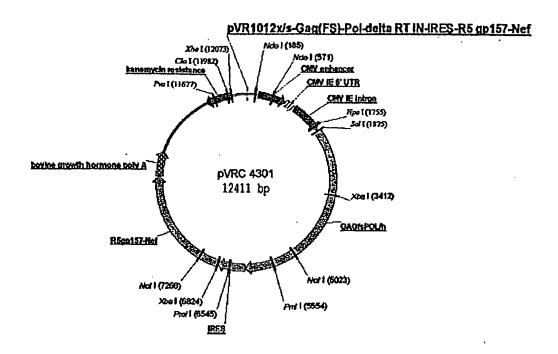


Figure 165

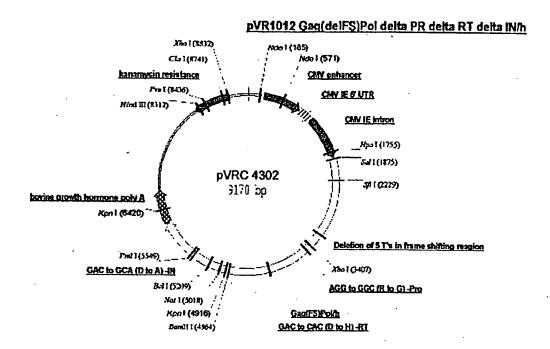


Figure 166

## pVR1012x/s SIV Gag(delFS)-Pol/h

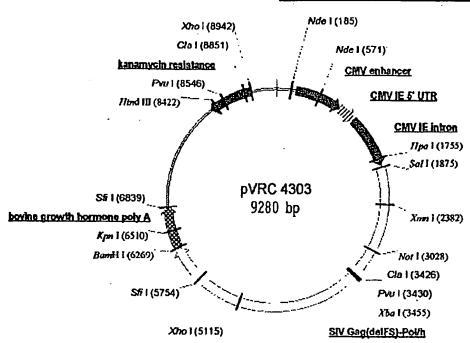


Figure 167

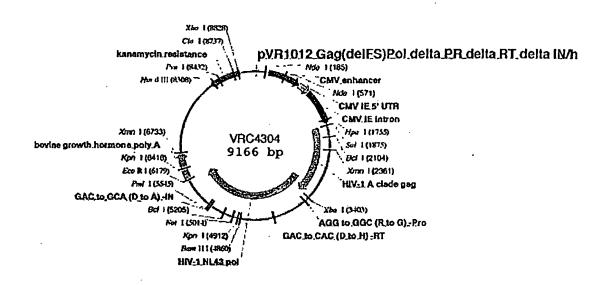


Figure 168

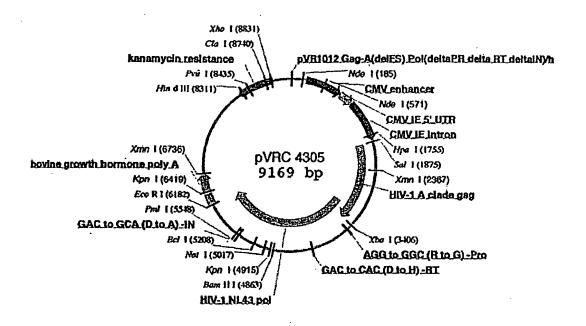


Figure 169

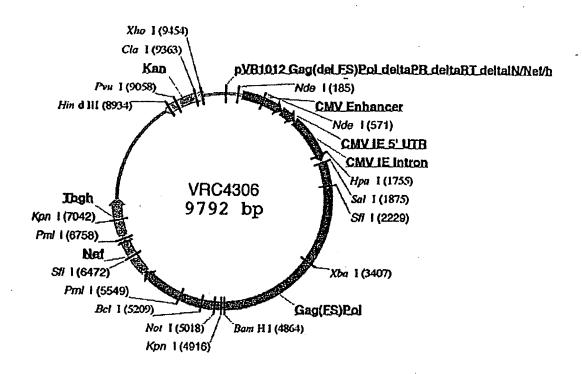


Figure 170

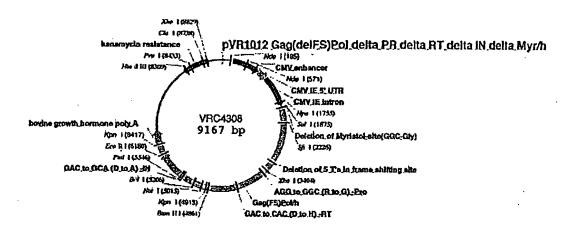


Figure 171

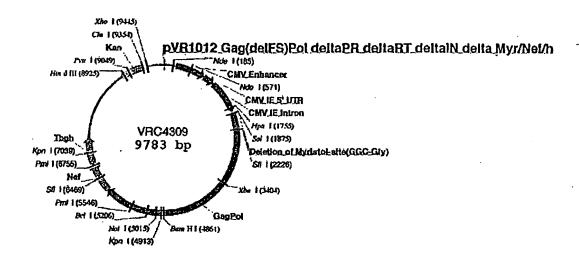


Figure 172

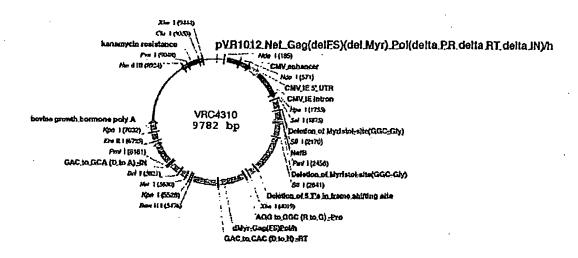


Figure 173

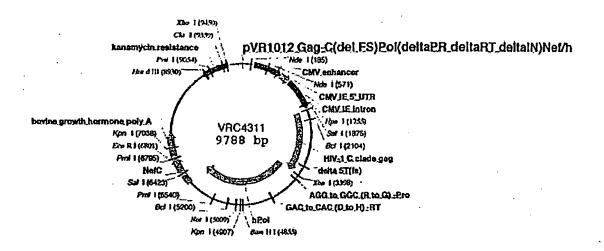


Figure 174

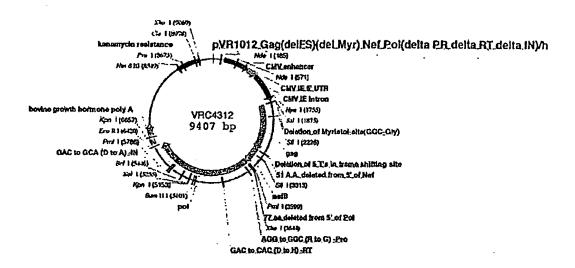


Figure 175

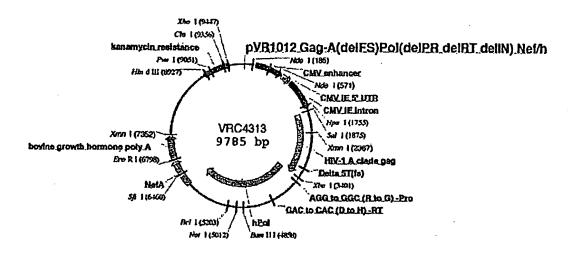


Figure 176

### SEQUENCE LISTING

<110> THE GOVERNMENT OF THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES NABLE, Gary J.
CHAKRABARTI, Bimal K.
HUANG, Yue

<120> MODIFICATIONS OF HIV Env, Gag, AND Pol ENHANCE IMMUNOGENICITY FOR GENETIC IMMUNIZATION

<130> NIH206.001VPC

<140> UNKNOWN

<141>

<150> US 60/225,097

<151> 2000-08-14

<150> US 60/252,115

<151> 2000-11-14

<150> US 60/279,257

<151> 2001-03-28

<160> 176

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 7547

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

### <400> 1

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtct gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgit gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatagga gttccgcgtt acataactta cggtaaatgg 420 eccgeetgge tgacegeeca acgaeeceeg eccattgacg teaataatga egtatgttee 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtitgtitt ggcaccaaaa tcaacgggac titccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960

| tagaagacac  | cgggaccgat | ccagcctccg | cggccgggaa | cggtgcattg | gaacgcggat | 1020 |
|-------------|------------|------------|------------|------------|------------|------|
| tccccgtgcc  | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca | cccctttggc | 1080 |
| tcttatgcat  | gctatactgt | ttttggcttg | gggcctatac | acccccgctt | ccttatgcta | 1140 |
| taggtgatgg  | tatagcttag | cctataggtg | tgggttattg | accattattg | accactcccc | 1200 |
| tattggtgac  | gatactttcc | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260 |
| tattggctat  | atgccaatac | tetgteette | agagactgac | acggactctg | tatttttaca | 1320 |
| ggatggggtc  | ccatttatta | tttacaaatt | cacatataca | acaacgccgt | ccccqtqcc  | 1380 |
| cgcagtttt   | attaaacata | gcgtgggatc | tccacqcqaa | tctcgggtac | gtgttccgga | 1440 |
| catgggctct  | teteeggtag | cggcggagct | tccacatcca | agccctggtc | ccatacctcc | 1500 |
| ageggeteat  | ggtcgctcgg | cageteettg | ctcctaacag | tggaggccag | acttaggcac | 1560 |
| agcacaatgc  | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | atatatatat | 1620 |
| gaaaatgagc  | gtggagattg | gactcacaca | gctgacgcag | atggaagact | taagggagg  | 1680 |
| gcagaagaag  | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
| attacaatac  | tottaacoot | agaagacaat | gtagtctgag | cagtactcgt | tactaccaca | 1800 |
| Cacaccacca  | gacataatag | chacagact  | aacadactdt | tcctttccat | gggtgttttg | 1860 |
| tgcagtcacc  | gtcgtcgaca | carataatca | gatatogogo | ccgctctaga | acaccators | 1000 |
| catasagaa   | aagtaccagc | acctatagea | ctagaactaa | cgctggggca | acaccatgeg | 1000 |
|             |            |            |            |            |            |      |
| acceptates  | acgaccegea | gegeeacega | gaagetgtgg | gtgaccgtgt | actacggegt | 2040 |
| caccacacata | cacaacctct | gggggaggg  | gccctgcgcc | agcgacgcca | aggeetaega | 2100 |
| gazastaata  | cacaacgcgc | tanagagaan | cyceryegeg | cccaccgacc | ccaaccccca | 2160 |
| ggaggraggra | gaggagatga | taaaaatata | cccgacacg  | tggaagaacg | acatggtgga | 2220 |
| geagatgeac  | tacatanaa  | taaaataaa  | ggaccagage | ctgaagccct | gegtgaaget | 2280 |
| caggagggg   | cgcgcgagcc | tgaagtgcac | cgacctgaag | aacgacacca | acaccaacag | 2340 |
| cagcagcggc  | cgcatgatta | tagagaaggg | cyagaccaag | aactgcagct | tcaacatcag | 2400 |
| caccagcate  | cgcggcaagg | cgcagaagga | graegeerre | ttctacaagc | tggacatcat | 2460 |
| ceceategae  | aacgacacca | ccagctacag | cctgaccage | tgcaacacca | gcgtgatcac | 2520 |
| ccaggeetge  | cccaaggtga | gettegagee | carccccaac | cactactgcg | cccccgccgg | 2580 |
| cttcaccatc  | ecgaagegea | aggacaagaa | gttcaacggc | aagggcccct | gcaccaacgt | 2640 |
| gageacegeg  | cagegeacce | acggcatccg | eccegtggtg | agcacccagc | tgctgctgaa | 2700 |
| cgaccgcctg  | geegaggagg | aggeggegae | ccgcagcgcc | aacttcgccg | acaacgccaa | 2760 |
| ggtgateate  | gegeageega | acgagagcgt | ggagatcaac | tgcacccgcc | ccaacaacaa | 2820 |
| caccegeaag  | agcatecaca | reggeeeegg | ccgcgccttc | tacaccaccg | gcgagatcat | 2880 |
| eggegaeate  | cgccaggccc | actgcaacct | gageegegee | aagtggaacg | acaccctgaa | 2940 |
| caagategeg  | accaagetge | gcgagcagtt | cggcaacaag | accatcgtgt | tcaagcacag | 3000 |
| cageggegge  | gaccccgaga | tegtgaecca | cagcttcaac | tgcggcggcg | agttcttcta | 3060 |
| ctgcaacagc  | acccagctgt | tcaacagcac | ctggttcaac | agcacctgga | gcaccgaggg | 3120 |
| cagcaacaac  | accgagggca | gcgacaccat | caccctgccc | tgccgcatca | agcagatcat | 3180 |
| caacatgtgg  | cagaaggtgg | gcaaggccat | gtacgccccc | cccatcagcg | gccagatccg | 3240 |
| ccgcagcage  | aacatcaccg | gcctgctgct | gacccgcgac | ggcggcaaca | gcaacaacga | 3300 |
| gagegagate  | ttccgcctgg | gcggcggcga | catgcgcgac | aactggcgca | gcgagctgta | 3360 |
| caagtacaag  | gtggtgaaga | tegageceet | gggcgtggcc | cccaccaagg | ccaagcgccg | 3420 |
| cgtggtgcag  | cgcgagaagc | gcgccgtggg | catcggcgcc | ctgttcctgg | gcttcctggg | 3480 |
| cgccgccggc  | agcaccatgg | gcgccgccag | catgaccctg | accgtgcagg | cccgccagct | 3540 |
| gctgagcggc  | atcgtgcagc | agcagaacaa | cctgctgcgc | gccatcgagg | cccagcagca | 3600 |
| cctgctgcag  | ctgaccgtgt | ggggcatcaa | gcagctgcag | gcccgcatcc | tggccgtgga | 3660 |
| gcgctacctg  | aaggaccagc | agctgctggg | catctggggc | tgcagcggca | agcttatctg | 3720 |
| caccaccgcc  | gtgccctgga | acgccagctg | gagcaacaag | agcctggagc | agatctggaa | 3780 |
| ccacaccacc  | tggatggagt | gggaccgcga | gatcaacaac | tacaccagcc | tgatccacag | 3840 |
| cctgatcgag  | gagagccaga | accagcagga | gaagaacgag | caggagctgc | tggagctgga | 3900 |
| caagtgggcc  | agcatgggcg | gcaagtggag | caagagcagc | gtgatcggct | ggcccgccgt | 3960 |
| gcgcgagcgc  | atgcgccgcg | ccgagcccgc | cgccgacggc | gtgggcgccg | tgagccgcga | 4020 |
| cctggagaag  | cacggcgcca | tcaccagcag | caacaccgcc | gccaacaacg | ccgcctgcgc | 4080 |
| ctggctggag  | gcccaggagg | aggaggaggt | gggcttcgca | gtgaccgcac | aggtggcact | 4140 |
| gcgcgcaatg  | acctacaagg | ccgccgtgga | cctgagccac | ttcctgaagg | agaagggcgg | 4200 |
| cctggagggc  | ctgatccaca | gccagcgccg | ccaggacatc | ctggacctgt | ggatctacca | 4260 |

```
cacccagggc tacttccccg actggcagaa ctacaccccc ggccccggcg tgcgctaccc 4320
cetgacette ggetggtget acaagetggt gecegtggag eeegacaagg tggaggagge 4380
caacaagggc gagaacacca gcctgctgca ccccgtgagc ctgcacggca tggcagcacc 4440
cgagcgcgag gtgctggagt ggcgcttcga cagccgcctg gccttccacc acgtggcccg 4500
cgagctgcac cccgagtact tcaagaactg ctgaggatcc agatctgctg tgccttctag 4560
ttgccagcca tctgttgttt gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac 4620
teceactgte ettteetaat aaaatgagga aattgeateg eattgtetga gtaggtgtea 4680
ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag 4740
caggcatgct ggggatgcgg tgggctctat gggtacccag gtgctgaaga attgacccgg 4800
ttcctcctgg gccagaaaga agcaggcaca tccccttctc tgtgacacac cctgtccacg 4860
eccetggttc ttagttccag ceccaetcat aggacactca tagetcagga gggetcegec 4920
ttcaatccca cccgctaaag tacttggage ggtctctccc tccctcatca gcccaccaaa 4980
ccaaacctag cctccaagag tgggaagaaa ttaaagcaag ataggctatt aagtgcagag 5040
ggagagaaaa tgcctccaac atgtgaggaa gtaatgagag aaatcataga attttaaggc 5100
catcatggcc ttaatcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc 5160
tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg 5220
ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg 5280
ccgcgttgct ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac 5340
gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg 5400
gaageteest egtgegetet cetgtteega ceetgeeget taceggatac etgteegeet 5460
ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg 5520
tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 5580
gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac 5640
tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt 5700
tettgaagtg gtggeetaac tacggetaca etagaagaac agtatttggt atetgegete 5760
tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 5820
ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 5880
ctcaagaaga teetttgate ttttctaegg ggtctgaege teagtggaac gaaaactcae 5940
gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 6000
aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc 6060
aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg 6120
cctgactcgg ggggggggg cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat 6180
accaggeetg aategeecca teatecagee agaaagtgag ggageeacgg ttgatgagag 6240
ctttgttgta ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg 6300
cgttgtcggg aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac 6360
aaagccgccg tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa 6420
ttctgattag aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt 6480
atcaatacca tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca 6540
gttccatagg atggcaagat cctggtatcg gtctgcgatt ccgactcgtc caacatcaat 6600
acaacctatt aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt 6660
gacgactgaa tooggtgaga atggcaaaag ottatgcatt totttocaga ottgttoaac 6720
aggecageca tracgetegt carcaaaate actegeatea accaaacegt tatteatteg 6780
tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg 6840
aatcgaatgc aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc 6900
aggatattct tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca 6960
tgcatcatca ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag 7020
ccagtttagt ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt 7080
cagaaacaac tetggegeat egggettece atacaatega tagattgteg cacetgattg 7140
cccgacatta tcgcgagccc atttataccc atataaatca gcatccatgt tggaatttaa 7200
togoggooto gagoaagaog titocogitg aatatggoto ataacaccoc tigitattact 7260
gtttatgtaa gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta 7320
acatcagaga ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta 7380
tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat 7440
aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat 7500
catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtc
                                                                  7547
```

<210> 2

```
<211> 8009
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegetet eggegatgae ggegaaaace teegacacat geageteeeg gagacqqtea 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg qaacqcqqat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080
tettatgeat getatactgt tittggettg gggeetatae acceegett cettatgeta 1140
taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560
agcacaatgo ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcqt tqctqccqcq 1800
cgcgccacca gacataatag ctgacagact aacagactgt teetttecat gggtetttte 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga acaccatgcg 1920
cgtgaaggag aagtaccage acctgtggcg ctggggctgg cgctggggca ccatgctgct 1980
gggcatgctg atgatctgca gcgccaccga gaagctgtgg gtgaccgtgt actacggcgt 2040
gcccgtgtgg aaggaggcca ccaccacct gctctgcgcc agcgacgcca aggcctacga 2100
caccgaggtg cacaacgtgt gggccaccca cgcctgcgtg cccaccgacc ccaacccca 2160
ggaggtggtg ctggtgaacg tgaccgagaa cttcgacatg tggaagaacg acatggtgga 2220
gcagatgcac gaggacatca tcagcctgtg ggaccagagc ctgaaqccct qcgtgaaqct 2280
gaccccctg tgcgtgagcc tgaagtgcac cgacctgaag aacgacacca acaccaacaq 2340
cagcagcggc cgcatgatca tggagaaggg cgagatcaag aactgcagct tcaacatcag 2400
caccagcate egeggeaagg tgeagaagga gtacgcette ttetacaage tggacateat 2460
ceccategae aacgacacea ecagetacag ectgaceage tgeaacacea gegtqateae 2520
ccaggcctgc cccaaggtga gcttcgagcc catccccaac cactactgcg cccccgccgg 2580
```

```
cttcaccatc ctgaagtgca aggacaagaa gttcaacggc aagggcccct gcaccaacgt 2640
gagcaccgtg cagtgcaccc acggcatccg ccccgtggtg agcacccagc tgctgctgaa 2700
cgaccgcctg gccgaggagg aggtggtgat ccgcagcgcc .aacttcgccg acaacgccaa 2760
ggtgatcatc gtgcagctga acgagagcgt ggagatcaac tgcacccgcc ccaacaacaa 2820
caccegcaag agcatecaca teggeecegg eegegeette tacaccaceg gegagateat 2880
cggcgacatc cgccaggccc actgcaacct gagccgcgcc aagtggaacg acaccctgaa 2940
caagatcgtg atcaagctgc gcgagcagtt cggcaacaag accatcgtgt tcaagcacag 3000
cageggegge gacecegaga tegtgaceca cagetteaac tgeggeggeg agttetteta 3060
ctgcaacagc acccagctgt tcaacagcac ctggttcaac agcacctgga gcaccgaggg 3120
cagcaacaac accgagggca gcgacaccat caccctgccc tgccgcatca agcagatcat 3180
caacatgtgg cagaaggtgg gcaaggccat gtacgcccc cccatcagcg gccagatccg 3240
ctgcagcagc aacatcaccg gcctgctgct gacccgcgac ggcggcaaca gcaacaacga 3300
gagcgagatc ttccgcctgg gcggcggcga catgcgcgac aactggcgca gcgagctgta 3360
caagtacaag gtggtgaaga tcgagcccct gggcgtggcc cccaccaagg ccaagcgccg 3420
cgtggtgcag cgcgagaagc gcgccgtggg catcggcgcc ctgttcctgg gcttcctggg 3480
egeegeegge ageaccatgg gegeegeeag catgaccetg accgtgeagg cccgccaget 3540
gctgagcggc atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca 3600
cctgctgcag ctgaccgtgt ggggcatcaa gcagctgcag gcccgcatcc tggccgtgga 3660
gegetacetg aaggaceage agetgetggg catetgggge tgeageggea agettatetg 3720
caccaccgcc gtgccctgga acgccagctg gagcaacaag agcctggagc agatctggaa 3780
ccacaccacc tggatggagt gggaccgcga gatcaacaac tacaccagcc tgatccacag 3840
cctgatcgag gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga 3900
caagtgggcc agcctgtgga actggttcaa catcaccaac tggctgtggt acatcaagct 3960
gttcatcatg atcgtgggcg gcctggtggg cctgcgcatc gtgttcgccg tgctgagcat 4020
cgtgaaccgc gtgcgccagg gctacagccc cctgagcttc cagacccacc tgcccatccc 4080
ccgcggcccc gaccgccccg agggcatcga ggaggagggc ggcgagcgcg accgcgaccg 4140
cacgateege etggtgaacg geageetgge eetgatetgg gacgacetge geageetgtg 4200
cctgttcagc taccaccgcc tgcgcgacct gctgctgatc gtgacccgca tcgtggagct 4260
getgggeege egeggetggg aggeeetgaa gtactggtgg aacetgetge agtactggag 4320
ccaggagetg aagaacageg ccgtgageet getgaacgee accgccatca tgggeggeaa 4380
gtggagcaag agcagcgtga tcggctggcc cgccgtgcgc gagcgcatgc gccgcgccga 4440
gcccgccgcc gacggcgtgg gcgccgtgag ccgcgacctg gagaagcacg gcgccatcac 4500
cagcagcaac accgccgcca acaacgccgc ctgcgcctgg ctggaggccc aggaggagga 4560
ggaggtgggc ttcgcagtga ccgcacaggt ggcactgcgc gcaatgacct acaaggccgc 4620
cgtggacctg agccacttcc tgaaggagaa gggcggcctg gagggcctga tccacagcca 4680
gegeegeeag gacateetgg acetgtggat ctaccacace cagggetact teecegactg 4740
gcagaactac acccceggee ceggegtgeg ctaccecetg accttegget ggtgctacaa 4800
getggtgeee gtggageeeg acaaggtgga ggaggeeaae aagggegaga acaeeageet 4860
getgeacece gtgageetge acggeatgge ageacecgag egegaggtge tggagtggeg 4920
cttcgacage cgcctggcct tccaccacgt ggcccgcgag ctgcaccccg agtacttcaa 4980
gaactgctga ggatccggat ccagatctgc tgtgccttct agttgccagc catctgttgt 5040
ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta 5100
ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc tggggggtgg 5160
ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc 5220
ggtgggctct atgggtaccc aggtgctgaa gaattgaccc ggttcctcct gggccaqaaa 5280
gaagcaggca catccccttc tetgtgacac accctgtcca cgcccctggt tcttagttcc 5340
agececacte ataggacact catageteag gagggeteeg cetteaatee caccegetaa 5400
agtacttgga gcggtctctc cctccctcat cagcccacca aaccaaacct agcctccaag 5460
agtgggaaga aattaaagca agataggcta ttaagtgcag agggagagaa aatgcctcca 5520
acatgtgagg aagtaatgag agaaatcata gaattttaag gccatcatgg ccttaatctt 5580
ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag 5640
ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 5700
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa qqccqcqttq ctqqcqtttt 5760
tecatagget eegeceeect gacgageate acaaaaateg acgeteaagt cagaggtgge 5820
gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct 5880
```

```
ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 5940
tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca 6000
agetgggetg tgtgcacgaa ceceeegtte agecegaceg etgegeetta teeggtaact 6060
atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta 6120
acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtggccta 6180
actacggeta cactagaaga acagtatttg gtatetgege tetgetgaag ceagttacet 6240
teggaaaaag agttggtage tettgateeg geaaacaaac cacegetggt ageggtggtt 6300
tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 6360
tettttetae ggggtetgae geteagtgga acgaaaacte acgttaaggg attttggtea 6420
tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat 6480
caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg 6540
cacctatete agegatetgt etatttegtt catecatagt tgeetgaete gggggggggg 6600
ggcgctgagg tctgcctcgt gaagaaggtg ttgctgactc ataccaggcc tgaatcgccc 6660
catcatccag ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc 6720
agttggtgat tttgaacttt tgctttgcca cggaacggtc tgcgttgtcg ggaagatgcg 6780
tgatctgatc cttcaactca gcaaaagttc gatttattca acaaagccgc cgtcccgtca 6840
agtcagcgta atgctctgcc agtgttacaa ccaattaacc aattctgatt agaaaaactc 6900
atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg 6960
aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag 7020
atcotggtat cggtctgcga ttccgactcg tccaacatca atacaaccta ttaatttccc 7080
ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga 7140
gaatggcaaa agcttatgca tttctttcca gacttgttca acaggccagc cattacgctc 7200
gtcatcaaaa tcactcgcat caaccaaacc gttattcatt cgtgattgcg cctgagcgag 7260
acgaaatacg cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg 7320
caggaacact gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac 7380
ctggaatgct gttttcccgg ggatcgcagt ggtgagtaac catgcatcat caggagtacg 7440
gataaaatgc ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat 7500
ctcatctgta acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc 7560
ategggette ecatacaate gatagattgt egeacetgat tgeeegacat tategegage 7620
ccatttatac ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga 7680
cgtttcccgt tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag 7740
ttttattgtt catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga 7800
cacaacgtgg ctttcccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 7860
gagcggatac atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt 7920
tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa 7980
aaataggcgt atcacgaggc cctttcgtc
```

```
<210> 3
<211> 7558
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 3
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagccg tcaggggcg tcagcgggtg tcggggtg cttaactatg cggcatcaga gcagattgta ctgagagtgc 120
ttggcggtt tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatggc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta ccggtaaatgg 420
```

```
ecegeetgge tgacegeeca acgaceceeg eccattgacg teaataatga egtatgttee 480
 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtacege ctatagaete tataggeaca eccetttgge 1080
tettatgeat getatactgt ttttggettg gggeetatac acceeegett cettatgeta 1140
taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaeteece 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100
accgaggtge acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagectgtgg gaccagagec tgaagecetg cgtgaagetg 2280
accccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
eccategaca acgacaceae cagetacage etgaceaget geaacaceag egtgateace 2520
caggeetgee ccaaggtgag ettegageee atecccaace actactgege eccegeegge 2580
ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggcccctg caccaacgtg 2640
agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760
accatcateg tgcagetgaa caccagegtg gagatcaaet gcaccegeee caacaacaae 2820
accegeaage gaattegeat eeagegegge eeeggeegeg cettegtgae categgeaag 2880
ateggeaaca tgegecagge ceaetgeaac ateageegeg ceaagtggaa caacaceetg 2940
aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000
cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc 3060
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagageg agatetteeg eetgggegge ggegacatge gegacaaetg gegeagegag 3360
ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggcccccac caaggccaag 3420
egeegegtgg tgeagegega gaagegegee gtgggeateg gegeeetgtt cetgggette 3480
ctgggcgccg ccggcagcac catgggcgcc gccagcatga ccctgaccgt gcaggcccgc 3540
cagetgetga geggeategt geageageag aacaacetge tgegegeeat egaggeeeag 3600
cagcacetge tgeagetgae egtgtgggge ateaageage tgeaggeeeg cateetggee 3660
gtggagcgct acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720
```

```
atctgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780
tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3840
cacagcetga tegaggagag ceagaaceag eaggagaaga acgageagga getgetggag 3900
ctggacaagt gggccagcat gggcggcaag tggagcaaga gcagcgtgat cggctggccc 3960
geogtgegeg agegeatgeg eegegeegag eeegeegeeg aeggegtggg egeegtgage 4020
egegacetgg agaageaegg egecateaee ageageaaea eegeegeeaa caacqeeqee 4080
tgcgcctggc tggaggccca ggaggaggag gaggtgggct tcgcagtgac cgcacaggtg 4140
gcactgcgcg caatgaccta caaggccgcc gtggacctga gccacttcct gaaggagaag 4200
ggcggcctgg agggcctgat ccacagccag cgccgccagg acatcctgga cctgtggatc 4260
taccacaccc agggetaett eccegaetgg cagaactaca ecceeggeee eggegtgege 4320
taccccctga cetteggetg gtgetacaag etggtgeeeg tggageeega caaggtggag 4380
gaggccaaca agggcgagaa caccagcetg etgcacceeg tgageetgca eggcatggca 4440
gcaccegage gegaggtget ggagtggege ttegacagee geetggeett ccaccaegtg 4500
gcccgcgagc tgcaccccga gtacttcaag aactgctgag gatccggatc cagatctgct 4560
gtgccttcta gttgccagcc atctgttgtt tgcccctccc ccgtgccttc cttgaccctg 4620
gaaggtgcca ctcccactgt cctttcctaa taaaatgagg aaattgcatc gcattgtctg 4680
agtaggtgtc attctattct ggggggtggg gtggggcagg acagcaaggg ggaggattgg 4740
gaagacaata gcaggcatgc tggggatgcg gtgggctcta tgggtaccca ggtgctgaag 4800
aattgacccg gttcctcctg ggccagaaag aagcaggcac atccccttct ctgtgacaca 4860
ccctgtccac gcccctggtt cttagttcca gccccactca taggacactc atagctcagg 4920
agggeteege etteaateee accegetaaa gtacttggag eggtetetee eteceteate 4980
ageceaceaa accaaaceta geeteeaaga gtgggaagaa attaaageaa gataggetat 5040
taagtgcaga gggagagaaa atgcctccaa catgtgagga agtaatgaga gaaatcatag 5100
aattttaagg ccatcatggc cttaatcttc cgcttcctcg ctcactgact cgctgcgctc 5160
ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac 5220
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 5280
ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 5340
caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc 5400
gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 5460
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 5520
teteagtteg gtgtaggteg ttegeteeaa getgggetgt gtgcacgaac cccccgttea 5580
gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga 5640
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 5700
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 5760
tatctgcgct ctgctgaagc cagttacctt cggaaaaaaga gttggtagct cttgatccgg 5820
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 5880
aaaaaaagga totcaagaag atootttgat ottttotacg gggtotgacg otcagtggaa 5940
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 6000
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 6060
tgacagttac caatgettaa teagtgagge acetatetea gegatetgte tatttegtte 6120
atccatagtt gcctgactcg ggggggggg gcgctgaggt ctgcctcgtg aagaaggtgt 6180
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 6240
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 6300
ggaacggtet gegttgtegg gaagatgegt gatetgatee tteaacteag caaaagtteg 6360
atttattcaa caaageegee gteeegteaa gteagegtaa tgetetgeea gtgttacaae 6420
caattaacca attotgatta gaaaaactca togagcatca aatgaaactg caatttatto 6480
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 6540
tcaccgaggc agttccatag gatggcaaga tcctggtatc ggtctgcgat tccgactcgt 6600
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 6660
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 6720
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 6780
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 6840
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 6900
teacetgaat caggatatte ttetaatace tggaatgetg ttttcccggg gategeagtg 6960
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 7020
```

```
aattccgtca gccagtttag tctgaccate tcatetgtaa catcattggc aacgctacct 7080 ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 7140 gcacctgatt gcccgacatt atcgcgagcc cattatacc catataaatc agcatccatg 7200 ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 7260 cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 7320 tggcaatgt aacatcagag attttgagac acaacgtggc tttcccccc cccccattat 7380 tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 7440 aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 7500 accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 7558
```

tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctet teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegecacea gacataatag etgacagaet aacagaetgt teettteeat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040

- 9 -

```
<210> 4
<211> 7351
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegettt eggegatgae ggegaaaace teegacacat geageteeeq qaqaeqqtea 60
cagettgtct gtaageggat geegggagea gacaageeeg teageggege teageggqtq 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
```

```
cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
accccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcageggee geatgateat ggagaaggge gagateaaga actgeagett caacateage 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520
caggeetgee ccaaggtgag ettegageee atecceaace actactgege eccegeegge 2580
ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggcccctg caccaacgtg 2640
agcacegtgc agtgcaccca eggcateege eeegtggtga gcacecaget getgetgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760
accatcatcg tgcagctgaa caccagcgtg gagatcaact gcacccgccc caacaacaac 2820
accegeaage gaattegeat ecagegegge eceggeegeg cettegtgae categgeaag 2880
ateggeaaca tgegeeagge ceaetgeaac ateageegeg ceaagtggaa caacaccetg 2940
aagcagateg acagcaagct gegegageag tteggeaaca acaagaceat catetteaag 3000
cagageageg geggegacee egagategtg acceacaget teaactgegg eggegagtte 3060
ttetaetgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atcogotgoa goagoaacat cacoggootg otgotgacoo gogacggogg caacagcaac 3300
aacgagageg agatetteeg eetgggegge ggegacatge gegacaactg gegeagegag 3360
ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggcccccac caaggccaag 3420
cgccgcgtgg tgcagcgcga gaagcgcgcc gtgggcatcg gcgccctgtt cctgggcttc 3480
ctgggcgccg ccggcagcac catgggcgcc gccagcatga ccctgaccgt gcaggcccgc 3540
cagetgetga geggeategt geageageag aacaacetge tgegegeeat egaggeeeag 3600
cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660
gtggagcgct acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720
ggtggcaagt ggtcaaaaag tagtgtgatt ggatggcctg ctgtaaggga aagaatgaga 3780
cgagctgagc cagcagcaga tggggtggga gcagtatctc gagacctaga aaaacatgga 3840
gcaatcacaa gtagcaatac agcagctaac aatgctgctt gtgcctggct agaagcacaa 3900
gaggaggaag aggtgggttt tccagtcaca cctcaggtac ctttaagacc aatgacttac 3960
aaggcagctg tagatcttag ccacttttta aaagaaaagg ggggactgga agggctaatt 4020
cactcccaaa gaagacaaga tatccttgat ctgtggatct accacacaca aggctacttc 4080
cctgattggc agaactacac accagggcca ggggtcagat atccactgac ctttggatgg 4140
tgctacaagc tagtaccagt tgagccagat aaggtagaag aggccaataa aggagagaac 4200
accagettgt tacaccetgt gageetgeat ggaatggatg accetgagag agaagtgtta 4260
gagtggaggt ttgacagccg cctagcattt catcacgtgg cccgagagct gcatccggag 4320
tacttcaaga actgctgagg atccagatot gctgtgcctt ctagttgcca gccatctgtt 4380
gtttgcccct cccccgtgcc ttccttgacc ctggaaggtg ccactccac tgtcctttcc 4440
taataaaatg aggaaattgc atcgcattgt ctgagtaggt gtcattctat tctggggggt 4500
ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggat 4560
gcggtgggct ctatgggtac ccaggtgctg aagaattgac ccggttcctc ctgggccaga 4620
aagaagcagg cacatcccct tetetgtgac acaccetgte cacgeceetg gttettagtt 4680
ccagececae teataggaea eteatagete aggagggete egeetteaat eccaeceget 4740
aaagtacttg gagcggtete teeeteecte atcageceae caaaccaaae etageeteea 4800
agagtgggaa gaaattaaag caagataggc tattaagtgc agagggagag aaaatgcctc 4860
caacatgtga ggaagtaatg agagaaatca tagaatttta aggccatcat ggccttaatc 4920
ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc 4980
ageteactea aaggeggtaa taeggttate cacagaatea ggggataaeg caggaaagaa 5040
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcqt tgctggcqtt 5100
tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg 5160
gegaaaceeg acaggactat aaagatacea ggegttteee eetggaaget ceetegtgeq 5220
ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag 5280
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc 5340
```

```
caagetggge tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa 5400
ctategtett gagtecaace eggtaagaca egacttateg ceaetggeag eagceaetgg 5460
taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc 5520
taactacggc tacactagaa gaacagtatt tggtatctgc gctctgctga agccagttac 5580
cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg 5640
tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 5700
gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt 5760
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagtttaa 5820
atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga 5880
ggcacctatc tcagcgatct gtctatttcg ttcatccata gttgcctgac tcgggggggg 5940
ggggcgctga ggtctgcctc gtgaagaagg tgttgctgac tcataccagg cctgaatcgc 6000
cccatcatcc agccagaaag tgagggagcc acggttgatg agagctttgt tgtagqtqqa 6060
ccagttggtg attttgaact tttgctttgc cacggaacgg tctgcgttgt cgggaagatg 6120
cgtgatctga tccttcaact cagcaaaagt tcgatttatt caacaaagcc gccgtcccgt 6180
caagtcagcg taatgctctg ccagtgttac aaccaattaa ccaattctga ttagaaaaac 6240
tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatttt 6300
tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca taggatggca 6360
agatectggt ateggtetge gatteegact egteeaacat caatacaace tattaattte 6420
ccctcgtcaa aaataaggtt atcaagtgag aaatcaccat gagtgacgac tgaatccggt 6480
gagaatggca aaagcttatg catttettte cagaettgtt caacaggeca gecattacge 6540
tegteateaa aateaetege ateaaceaaa eegttattea ttegtgattg egeetgageg 6600
agacgaaata cgcgatcgct gttaaaagga caattacaaa caggaatcga atgcaaccgg 6660
cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata ttcttctaat 6720
acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcatc atcaggagta 6780
cggataaaat gcttgatggt cggaagaggc ataaattccg tcagccagtt tagtctgacc 6840
atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa caactctggc 6900
gcatcgggct tcccatacaa tcgatagatt gtcgcacctg attgcccgac attatcgcga 6960
gcccatttat acccatataa atcagcatcc atgttggaat ttaatcgcgg cctcgagcaa 7020
gacgtttccc gttgaatatg gctcataaca ccccttgtat tactgtttat gtaagcagac 7080
agttttattg ttcatgatga tatattttta tcttgtgcaa tgtaacatca gagattttga 7140
gacacaacgt ggctttcccc cccccccat tattgaagca tttatcaggg ttattgtctc 7200
atgageggat acatatttga atgtatttag aaaaataaac aaataggggt teegegeaca 7260
tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat 7320
aaaaataggc gtatcacgag gccctttcgt c
                                                                  7351
```

```
<210> 5
<211> 8008
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 5
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccggagca gacaagccg tcagggegg tcagcgggtg tcgggggtg tcggggtg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgat attgactagt tattaatagt aatcaattac 360
ggggtcatta gtcatagcc catatatgga gttccgcgtt accataattga cggtaagtg 420
cccgcctggc tgaccgcca acgaccccg cccattgacg gtggagtatt tacggtaaac 540
ccatagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
```

- 11 -

tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acgggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teecegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccqtqcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegecacca gacataatag etgacagaet aacagaetgt teetttecat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatqcqc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040 cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccag 2160 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtqqaq 2220 cagatgcacg aggacatcat cagcctgtgg gaccagagec tgaagecetg cgtgaagetg 2280 accecectgt gegtgageet gaagtgeace gaeetgaaga aegacaceaa caecaacage 2340 agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400 accagcatee geggeaaggt geagaaggag tacgcettet tetacaaget ggacateate 2460 cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cqtgatcacc 2520 caggeetgee ccaaggtgag ettegageee atecceaace actactgege eccegeegge 2580 ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggcccctg caccaacgtg 2640 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700 ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacqccaaq 2760 accatcatcg tgcagctgaa caccagcgtg gagatcaact gcacccgccc caacaacaac 2820 accegeaage gaattegeat ceagegegge eeeggeegeg eettegtgae eateggeaag 2880 ateggeaaca tgegeeagge ceaetgeaac ateageegeg ceaagtggaa caacaccetq 2940 aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000 cagagcagcg gcggcgaccc cgagatcgtg acccacaget tcaactgcgg cggcgagttc 3060 ttetactgea acageaceca getgtteaac ageacetggt teaacageac etggageace 3120 gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180 atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240 atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300 aacgagagcg agatetteeg eetgggegge ggegacatge gegacaactg gegeagegag 3360 ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggcccccac caaggccaag 3420 cgccgcgtgg tgcagcgcga gaagcgcgcc gtgggcatcg gcgccctgtt cctgggcttc 3480 ctgggegeeg ceggeageac catgggegee gecageatga ceetgacegt geaggeeege 3540 cagetgetga geggeategt geageageag aacaacetge tgegegeeat egaggeeeag 3600 cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctqqcc 3660 gtggageget acetgaagga eeageagetg etgggeatet ggggetgeag eggeaagett 3720 atotgoacca cogoogtgoo otggaacgoo agotggagoa acaagagoot ggagoaqato 3780 tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3840

| cacagcctga | tcgaggagag | ccagaaccag | caggagaaga | acgagcagga | gctgctggag | 3900 |
|------------|------------|------------|------------|------------|------------|------|
| ctggacaagt | gggccagcct | gtggaactgg | ttcaacatca | ccaactggct | gtggtacatc | 3960 |
| aagctgttca | tcatgatcgt | gggcggcctg | gtgggcctgc | gcatcgtgtt | cgccgtgctg | 4020 |
|            |            |            |            |            | ccacctgccc |      |
| atcccccgcg | gccccgaccg | ccccgagggc | atcgaggagg | agggcggcga | gcgcgaccgc | 4140 |
| gaccgcacga | tccgcctggt | gaacggcagc | ctggccctga | tctgggacga | cctgcgcagc | 4200 |
|            |            |            |            |            | ccgcatcgtg |      |
|            |            |            |            |            | gctgcagtac |      |
|            |            |            |            |            | catcatgggc |      |
|            |            |            |            |            | catgcgccgc |      |
|            |            |            |            |            | gcacggcgcc |      |
|            |            |            |            |            | ggcccaggag |      |
|            |            |            |            |            | gacctacaag |      |
|            |            |            |            |            | cctgatccac |      |
|            |            |            |            |            | ctacttcccc |      |
|            |            |            |            |            | cggctggtgc |      |
|            |            |            |            |            | cgagaacacc |      |
|            |            |            |            |            | ggtgctggag |      |
|            |            |            |            |            | ccccgagtac |      |
|            |            |            |            |            | atctgttgtt |      |
|            |            |            |            |            | cctttcctaa |      |
|            |            |            |            |            | ggggggtggg |      |
|            |            |            |            |            | tggggatgcg |      |
|            |            |            |            |            | ggccagaaag |      |
|            |            |            |            |            | cttagttcca |      |
|            |            |            |            |            | acccgctaaa |      |
|            |            |            |            |            | gcctccaaga |      |
|            |            |            |            |            | atgcctccaa |      |
|            |            |            |            |            | cttaatcttc |      |
|            |            |            |            |            | cggtatcagc |      |
|            |            |            |            |            | gaaagaacat |      |
|            |            |            |            |            | tggcgttttt |      |
|            |            |            |            |            | agaggtggcg |      |
| aaacccgaca | ggactataaa | gataccaggc | gtttccccct | ggaagctccc | tegtgegete | 5880 |
|            |            |            |            |            | cgggaagcgt |      |
| ggcgctttct | catageteae | gctgtaggta | tctcagttcg | gtgtaggtcg | ttcgctccaa | 6000 |
|            |            |            |            |            | ccggtaacta |      |
| tcgtcttgag | tccaacccgg | taagacacga | cttatcgcca | ctggcagcag | ccactggtaa | 6120 |
| caggattagc | agagcgaggt | atgtaggcgg | tgctacagag | ttcttgaagt | ggtggcctaa | 6180 |
| ctacggctac | actagaagaa | cagtatttgg | tatctgcgct | ctgctgaagc | cagttacctt | 6240 |
|            |            |            |            |            | gcggtggttt |      |
| ttttgtttgc | aagcagcaga | ttacgcgcag | aaaaaagga  | tctcaagaag | atcctttgat | 6360 |
|            |            |            |            |            | ttttggtcat |      |
|            |            |            |            |            | gttttaaatc |      |
|            |            |            |            |            | teagtgagge |      |
| acctatctca | gcgatctgtc | tatttcgttc | atccatagtt | gcctgactcg | 999999999  | 6600 |
| gcgctgaggt | ctgcctcgtg | aagaaggtgt | tgctgactca | taccaggcct | gaatcgcccc | 6660 |
| atcatccagc | cagaaagtga | gggagccacg | gttgatgaga | gctttgttgt | aggtggacca | 6720 |
| gttggtgatt | ttgaactttt | gctttgccac | ggaacggtct | gcgttgtcgg | gaagatgcgt | 6780 |
| gatctgatcc | ttcaactcag | caaaagttcg | atttattcaa | caaagccgcc | gtcccgtcaa | 6840 |
| gtcagcgtaa | tgctctgcca | gtgttacaac | caattaacca | attctgatta | gaaaaactca | 6900 |
|            |            |            |            |            | atatttttga |      |
|            |            |            |            |            | gatggcaaga |      |
|            |            |            |            |            | taatttcccc |      |
| tcgtcaaaaa | taaggttatc | aagtgagaaa | tcaccatgag | tgacgactga | atccggtgag | 7140 |

```
aatggcaaaa gcttatgcat ttctttcag acttgttcaa caggccagcc attacgctcg 7200 tcatcaaaat cactcgcatc aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc 7320 aggaacactg ccagcgcatc aacaatattt tcacctgaat caggatattc ttctaatacc 7380 tggaatgctg ttgatggcgg aagggcata aattccgtca gccagtttag tctgaccatc 7500 tcatctgtaa catcattggc aacgctacct ttgccatgtt tcagaaacaa ctctggcgca 7560 tcgggcttcc catacaatcg atagattgtc gcacctgatt gcccgacatt acggagcatc acgcggccatc acgggcttcc catacaatcg atagattgtc gcacctgatt gcccgacatt acgcgagcc 7620 cattatacc cataataaatc agcatccatg ttggaattta acgcggcct cgagcaagac 7680 gtttcccgtt gaatatggct catacaaccc cttgtattac tggttatga accatcagg 7800 acaacgtggc tttcccccc cccccattat tggaagcatt acggggttcc gcgcacattt 7860 agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt 7920 ccccgaaaag tgccacctga cgtctaagaa accattatta tcatggact aacctataaa 7980 aataggcgta tcacgaggcc ctttcgtc
```

```
<210> 6
<211> 6970
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
```

#### <400> 6

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaaccgte agategeetg gagacgccat ceacqetqtt ttqacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcqqqtac qtqttccqqa 1440 catgggetet tetecggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

```
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
 głgaaggaga agłaccagca ccłgłggcgc tggggcłggc gcłggggcac całgcłgcłg 1980
 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
 cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 2100
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
 cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaagetg 2280
 accecetgt gegtgageet gaagtgeace gaeetgaaga acgaeaceaa caccaacage 2340
 agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
 accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
 cccatcgaca acgacaccac cagctacage ctgaccaget gcaacaccag cgtgatcacc 2520
 caggectgee ccaaggtgag ettegagece atceccatee actactgege eccegecgge 2580
 ttegecatee tgaagtgeaa caacaagace tteaacggea eeggeeeetg caccaacgtg 2640
 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700
 ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760
 accatcatcg tgcagctgaa caccagcgtg gagatcaact gcacccgccc caacaacaac 2820
 accegeaage gaattegeat ecagegegge eceggeegeg cettegtgae categgeaag 2880
 ateggeaaca tgcgccagge ccactgeaac ateageegeg ccaagtggaa caacaccetg 2940
 aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000
 cagageageg geggegaece egagategtg acceaeaget teaactgegg eggegagtte 3060
 ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
 gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180
 atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
 atcegetgea geageaacat caceggeetg etgetgacee gegaeggegg caacageaac 3300
 aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360
 ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggcccccac caaggccaag 3420
 egecgegtgg tgeagegega gaagegegee gtgggeateg gegeeetgtt cetgggette 3480
 ctgggegeeg ceggeageac catgggegee gecageatga ceetgacegt geaggeeege 3540
 cagctgctga gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag 3600
 cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660
 gtggagcgct acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720
 atctgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780
 tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3840
 cacageetga tegaggagag eeagaaceag caggagaaga aegageagga getgetggag 3900
 ctggacaagt gggccagcct gtggaactgg ttcaacatca ccaactggct gtggtgagga 3960
 tecagatetg etgtgeette tagttgeeag ceatetgttg tttgeeeete eccegtgeet 4020
 tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca 4080
 tegeattgte tgagtaggtg teattetatt etggggggtg gggtggggea ggacageaag 4140
ggggaggatt gggaagacaa tagcaggcat gctggggatg cggtgggctc tatgggtacc 4200
caggtgctga agaattgacc cggttcctcc tgggccagaa agaagcaggc acatcccctt 4260
ctctgtgaca caccetgtee acgeeectgg ttcttagtte cageeceact cataggacae 4320
tcatagetca ggagggetce gcettcaate ccaccegeta aagtacttgg ageggtetet 4380
ccctcctca tcagcccacc aaaccaaacc tagcctccaa gagtgggaag aaattaaagc 4440
aagataggct attaagtgca gagggagaga aaatgcctcc aacatgtgag gaagtaatga 4500
gagaaatcat agaattttaa ggccatcatg gccttaatct tccgcttcct cgctcactga 4560
ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat 4620
acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca 4680
aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgccccc 4740
tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata 4800
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc 4860
gettacegga tacetgteeg cettteteee ttegggaage gtggegettt etcatagete 4920
acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 4980
```

```
acceccegtt cagecegace getgegeett ateeggtaae tategtettg agtecaacee 5040
ggtaagacac gacttatege cactggcage agecactggt aacaggatta gcagagegag 5100
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 5160
aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 5220
ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca 5280
gattacgcgc agaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 5340
cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 5400
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 5460
gtaaacttgg tetgacagtt accaatgett aatcagtgag geacetatet cagegatetg 5520
tetatttegt teatecatag ttgeetgaet eggggggggg gggegetgag gtetgeeteg 5580
tgaagaaggt gttgctgact cataccaggc ctgaatcgcc ccatcatcca gccagaaagt 5640
gagggagcca cggttgatga gagctttgtt gtaggtggac cagttggtga ttttgaactt 5700
ttgctttgcc acggaacggt ctgcgttgtc gggaagatgc gtgatctgat ccttcaactc 5760
cagtgttaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5880
tgcaatttat tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat 5940
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg 6000
attecgacte gtecaacate aatacaacet attaatttee cetegteaaa aataaggtta 6060
tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaatggcaa aagcttatgc 6120
atttctttcc agacttgttc aacaggccag ccattacgct cgtcatcaaa atcactcgca 6180
tcaaccaaac cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg 6240
ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 6300
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg 6360
gggatcgcag tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 6420
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6480
gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 6540
cgatagattg tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa 6600
tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg 6660
ctcataacac cccttgtatt actgtttatg taagcagaca gttttattgt tcatgatgat 6720
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttccccc 6780
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6840
tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct 6900
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6960
ccctttcgtc
```

```
<210> 7
<211> 6787
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 7
tcgcgcgttt cggtgatgac ggtgaaaacc cagcttgtc gfaaagcggat gccggagca gacaagccg tcaggggcgg tcaggggtg tcggggtg cttaactatg cggcatcaga gcagattgta ctgagagtg l20 tcgggggtg tcggggatga ccgacagat gcgtaagag gcgaattgta ctgagagtg l80 accatatgcg gtgtgaaata ccgcacagat tcatatggca ttggcatcata tcatatata tcgccatgtt gacattgat gggtcatta ggtcatata ccgccatgtt gacattgat ggtcatata ggtcattat ggggtcatta gtcataagca catatatgga gtccgcggt tcataatatg tacattata ttggctcatg 300 tcataagtac cccgcctggc tgaccgcca acgaccccg cccattgacg tcaataatga ccgtaagtgc tgaccgcca acgaccccg tcataatatg tacattata ttggctcatg 300 tcataagtaac ccaatagga ctttccattg acgtcaatgg gtggagtatt tacaggtaaac 540
```

```
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatage ggtttgacte acggggattt ccaagtetee accecattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcqtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegitta gigaacegic agategeeig gagaegeeat ceaegeight tigaceteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacqcqqat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eccetttgge 1080
tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
accecectgt gegtgageet gaagtgeace gaeetgaaga acgaeaceaa caccaacage 2340
agcageggee geatgateat ggagaaggge gagateaaga actgeagett caacateage 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520
caggeetgee ccaaggtgag ettegageee atceccaace actaetgege eccegeegge 2580
ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggcccctg caccaacgtg 2640
agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760
accatcateg tgcagetgaa caccagegtg gagatcaact gcaccegece caacaacaac 2820
accegeaage gaattegeat eeagegegge eeeggeegeg eettegtgae categgeaag 2880
ateggeaaca tgegeeagge ceaetgeaac ateageegeg ceaagtggaa caacaccetq 2940
aagcagateg acagcaaget gegegageag tteggeaaca acaagaceat catetteaag 3000
cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc 3060
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360
ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggccccac caaggccaag 3420
cttaccgtgc aggcccgcca gctgctgagc ggcatcgtgc agcagcagaa caacctgctg 3480
egegecateg aggeceagea geacetgetg cagetgaceg tgtggggcat caageagetg 3540
caggecegea tectggeegt ggagegetae etgaaggace ageagetgga geagatetqq 3600
aaccacacca cctggatgga gtgggaccgc gagatcaaca actacaccag cctgatccac 3660
aqcctgatcg aggagagcca gaaccagcag gagaagaacg agcaggagct gctggagctg 3720
qacaagtggg ccagcetgtg gaactggttc aacatcacca actggctgtg gtgaggatcc 3780
agatetgetg tgeettetag ttgecageea tetgttgttt geeceteece egtgeettee 3840
```

|            |              |              |                              |              | 101/0        | JU112J1. |
|------------|--------------|--------------|------------------------------|--------------|--------------|----------|
| ttgaccctgg | g aaggtgccac | : teccaetgte | ctttcctaat                   | : aaaatgagga | aattgcatcg   | 3900     |
| cattgtctga | a gtaggtgtca | ttctattctg   | , gggggtggg                  | r tqqqqcaqqa | caqcaaqqqq   | 3960     |
| gaggattggg | , aagacaatag | r caggcatgct | : ggggatgcgg                 | , tgggctctat | qqqtacccaq   | 4020     |
| gtgctgaaga | attgacccgg   | ttcctcctgg   | gccagaaaga                   | agcaggcaca   | teceettete   | 4080     |
| tgtgacacac | : cctgtccacg | cccctggttc   | : ttagttccag                 | ccccactcat   | aggacactca   | 4140     |
| tageteagga | ı gggctccgcc | ttcaatccca   | cccgctaaag                   | tacttggage   | ggtctctccc   | 4200     |
| teceteatea | gcccaccaaa   | ccaaacctag   | r cctccaagag                 | tgggaagaaa   | ttaaagcaag   | 4260     |
| ataggctatt | : aagtgcagag | ggagagaaaa   | tgcctccaac                   | atgtgaggaa   | gtaatgagag   | 4320     |
| aaatcataga | ı attttaaggc | catcatggco   | : ttaatcttcc                 | getteetege   | tcactgactc   | 4380     |
| gctgcgctcg | gtcgttcggc   | tgcggcgagc   | ggtatcagct                   | cactcaaagg   | cggtaatacg   | 4440     |
| gttatccaca | gaatcagggg   | ataacgcagg   | aaagaacatg                   | tgagcaaaag   | qccaqcaaaa   | 4500     |
| ggccaggaac | : cgtaaaaagg | ccgcgttgct   | ggcgttttc                    | cataggetee   | gccccctga    | 4560     |
| cgagcatcac | : aaaaatcgac | gctcaagtca   | gaggtggcga                   | aacccgacag   | gactataaag   | 4620     |
| ataccaggcg | tttccccctg   | gaagctccct   | <ul><li>cgtgcgctct</li></ul> | cctgttccga   | ccctaccact   | 4680     |
| taccggatac | ctgtccgcct   | ttctcccttc   | gggaagcgtg                   | gcgctttctc   | atageteacg   | 4740     |
| ctgtaggtat | ctcagttcgg   | tgtaggtcgt   | tegetecaag                   | ctgggctgtg   | tocacgaacc   | 4800     |
| ccccgttcag | cccgaccgct   | gcgccttatc   | cggtaactat                   | cqtcttqaqt   | ccaacccoot   | 4860     |
| aagacacgac | ttatcgccac   | tggcagcagc   | cactggtaac                   | aggattagca   | gagcgaggta   | 4920     |
| rgtaggcggt | gctacagagt   | tcttgaagtg   | gtggcctaac                   | tacggctaca   | ctagaagaac   | 4980     |
| agtatttggt | atctgcgctc   | tgctgaagcc   | agttaccttc                   | ggaaaaaqaq   | ttagtagete   | 5040     |
| ttgatccggc | aaacaaacca   | ccgctggtag   | cggtggtttt                   | tttqtttqca   | agcagcagat   | 5100     |
| cacgcgcaga | aaaaaaggat   | ctcaagaaga   | tcctttgatc                   | ttttctacqq   | ggtctgacgc   | 5160     |
| ccagtggaac | gaaaactcac   | gttaagggat   | tttggtcatg                   | agattatcaa   | aaaggatett   | 5220     |
| cacctagatc | cttttaaatt   | aaaaatgaag   | ttttaaatca                   | atctaaaqta   | tatatgagta   | 5280     |
| aacttggtct | gacagttacc   | aatgcttaat   | cagtgaggca                   | cctatctcag   | cgatctgtct   | 5340     |
| atttcgttca | tccatagttg   | cctgactcgg   | 999999999                    | cgctgaggtc   | tacctcataa   | 5400     |
| agaaggtgtt | gctgactcat   | accaggcctg   | aatcgcccca                   | tcatccagcc   | agaaagtgag   | 5460     |
| ggagccacgg | ttgatgagag   | ctttgttgta   | ggtggaccag                   | ttggtgattt   | tgaacttttg   | 5520     |
| ctttgccacg | gaacggtctg   | cgttgtcggg   | aagatgcgtg                   | atctgatcct   | tcaactcage   | 5580     |
| aaaagttcga | tttattcaac   | aaagccgccg   | tcccgtcaag                   | tcagcgtaat   | gctctgccag   | 5640     |
| tgttacaacc | aattaaccaa   | ttctgattag   | aaaaactcat                   | cqaqcatcaa   | atgaaactgc   | 5700     |
| aacccaccca | tatcaggatt   | atcaatacca   | tatttttgaa                   | aaaqccqttt   | ctgtaatgaa   | 5760     |
| ggagaaaact | caccgaggca   | gttccatagg   | atggcaagat                   | cctqqtatcq   | gtctgcgatt   | 5820     |
| ccgactcgtc | caacatcaat   | acaacctatt   | aatttcccct                   | cgtcaaaaat   | aaggttatca   | 5880     |
| agtgagaaat | caccatgagt   | gacgactgaa   | tccggtgaga                   | atggcaaaag   | cttatgcatt   | 5940     |
| tetttecaga | cttgttcaac   | aggccagcca   | ttacgctcgt                   | catcaaaatc   | actogoatoa   | 6000     |
| accaaaccgt | tattcattcg   | tgattgcgcc   | tgagcgagac                   | gaaatacgcg   | atcactatta   | 6060     |
| aaaggacaat | tacaaacagg   | aatcgaatgc   | aaccggcgca                   | ggaacactgc   | caqcqcatca   | 6120     |
| acaatatttt | cacctgaatc   | aggatattct   | tctaatacct                   | ggaatgctgt   | tttcccaaaa   | 6180     |
| atcgcagtgg | tgagtaacca   | tgcatcatca   | ggagtacgga                   | taaaatqctt   | gatggtcgga   | 6240     |
| agaggcataa | attccgtcag   | ccagtttagt   | ctgaccatct                   | catctgtaac   | atcattogca   | 6300     |
| acgctacctt | tgccatgttt   | cagaaacaac   | tctggcgcat                   | cagacttccc   | atacaatcga   | 6360     |
| tagattgtcg | cacctgattg   | cccgacatta   | tcgcgagece                   | atttataccc   | atataaatca . | 6420     |
| gcatccatgt | tggaatttaa   | tegeggeete   | gagcaagacg                   | tttcccgttg   | aatatggctc   | 6480     |
| ataacacccc | ttgtattact   | gtttatgtaa   | gcagacagtt                   | ttattqttca   | tgatgatata   | 6540     |
| tttttatctt | gtgcaatgta   | acatcagaga   | ttttgagaca                   | caacqtqqct   | ttccccccc    | 6600     |
| ccccattatt | gaagcattta   | tcagggttat   | tgtctcatga                   | gcggatacat   | atttqaatqt   | 6660 -   |
| atttagaaaa | ataaacaaat   | aggggttccg   | cgcacatttc                   | cccgaaaagt   | accacctasc   | 6720     |
| gtctaagaaa | ccattattat   | catgacatta   | acctataaaa                   | ataggcgtat   | cacgaggccc   | 6780     |
| tttcgtc    |              |              | •                            |              |              | 6787     |
|            |              |              |                              |              | •            |          |

<210> 8 <211> 6607

<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 8 tegegegttt eggtgatgac ggtgaaaace tetgacacat geageteeeg gagacggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgacg taagtacege ctatagacte tataggeaca eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagaetgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040 cccgtgtgga aggaggccac caccacctg ctctgcgcca gcgacgccaa ggcctacgac 2100 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280 acccccetgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340 agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400 accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460 cccatcgaca acgacaccac cagctacage etgaccaget gcaacaccag egtgatcace 2520 caggoetgee ccaaggtgag ettegageee atccccaace actactgege eccegeegge 2580 ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggcccctg caccaacgtg 2640 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700 ggcagcetgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760 accatcateg tgcagetgaa caccagegtg gagatcaact gcaccegeee caacaacaac 2820 accegeaage gaattegeat ceagegegge eeeggeegeg cettegtgac categgeaag 2880

```
atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940
aagcagateg acagcaaget gegegageag tteggeaaca acaagaceat catetteaag 3000
cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc 3060
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360
ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggccccac caaggccaag 3420
cttaccgtgc aggcccgcca gctgctgagc ggcatcgtgc agcagcagaa caacctgctg 3480
cgcgccatcg aggcccagca gcacctgctg cagctgaccg tgtggggcat caagcagctg 3540
caggeeegea teetggeegt ggagegetac etgaaggace ageagetega gtgaggatee 3600
agatetgetg tgccttctag ttgccageca tetgttgttt geceeteece egtgeettee 3660
ttgaccetgg aaggtgccac teceactgte ettteetaat aaaatgagga aattgcateg 3720
cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg 3780
gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat gggtacccag 3840
gtgctgaaga attgacccgg ttcctcctgg gccagaaaga agcaggcaca tccccttctc 3900
tgtgacacac cctgtccacg cccctggttc ttagttccag ccccactcat aggacactca 3960
tageteagga gggeteegee tteaateeea eeegetaaag taettggage ggteteteee 4020
tccctcatca gcccaccaaa ccaaacctag cctccaagag tgggaagaaa ttaaagcaag 4080
ataggctatt aagtgcagag ggagagaaaa tgcctccaac atgtgaggaa gtaatgagag 4140
aaatcataga attttaaggc catcatggcc ttaatcttcc gcttcctcgc tcactgactc 4200
getgegeteg gtegttegge tgeggegage ggtateaget cacteaaagg eggtaataeg 4260
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaa gccagcaaaa 4320
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 4380
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 4440
ataccaggeg tttccccctg gaageteect egtgegetet cetgtteega eeetgeeget 4500
taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 4560
ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 4620
ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 4680
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 4740
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4800
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 4860
ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 4920
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 4980
tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 5040
cacctagate ettttaaatt aaaaatgaag ttttaaatea atetaaagta tatatgagta 5100
aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5160
agaaggtgtt gctgactcat accaggcctg aatcgcccca tcatccagcc agaaagtgag 5280
ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt tgaacttttg 5340
ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatcct tcaactcaqc 5400
aaaagttcga tttattcaac aaagccgccg tcccgtcaag tcagcgtaat gctctgccag 5460
tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5520
aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa 5580
ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt 5640
cegactegte caacateaat acaacetatt aattteeect egteaaaaat aaggttatea 5700
agtgagaaat caccatgagt gacgactgaa teeggtgaga atggcaaaag ettatgcatt 5760
tetttecaga ettgtteaac aggecageca ttacgetegt cateaaaate actegeatea 5820
accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 5880
aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 5940
acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 6000
atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaatgctt gatggtcgga 6060
agaggcataa attoogtoag coagtttagt otgacoatot catotgtaac atcattggca 6120
acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6180
```

tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca 6240

```
gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg aatatggctc 6300
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6360
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct ttccccccc 6420
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6480
atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6540
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6600
tttcqtc
                                                                   6607
<210> 9
<211> 7042
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 9
tegegegtht eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat eeaegetgtt ttgaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080
tettatgeat getatactgt tittggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaeteece 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetch teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560
agcacaatge ceaceaceae cagtgtgeeg cacaaggeeg tggeggtagg gtatgtgtet 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100
```

```
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcetgtgg gaccagagee tgaageeetg egtgaagetg 2280
accecetgt gegtgageet gaagtgeace gaeetgaaga acgaeaceaa caccaacage 2340
agcageggee geatgateat ggagaaggge gagateaaga actgeagett caacateage 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
eccategaca acgacaceae cagetacage etgaceaget geaacaceag egtgateace 2520
caggeotice ccaaggigag ettegagece atececaace actactgege eccegeegge 2580
ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggcccctg caccaacgtg 2640
agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760
accatcateg tgcagetgaa caccagegtg gagatcaact gcaccegeee caacaacaac 2820
accegeaage gaattegeat ecagegege eceggeegeg cettegtgae categgeaag 2880
ateggeaaca tgegeeagge ceaetgeaac ateageegeg ceaagtggaa caacaceetg 2940
aagcagateg acageaaget gegegageag tteggeaaca acaagaceat catetteaag 3000
cagageageg geggegaeee egagategtg acceaeaget teaactgegg eggegagtte 3060
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagagcg agatetteeg eetgggegge ggegacatge gegacaaetg gegeagegag 3360
ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggcccccac caaggccaag 3420
cgccgcgtgg tgcagcgcga gaagcgcgcc gtgggcatcg gcgccctgtt cctgggcttc 3480
ctgggcgccg ccggcagcac catgggcgcc gccagcatga ccctgaccgt gcaggcccgc 3540
cagetgetga geggeategt geageageag aacaacetge tgegegeeat egaggeeeag 3600
cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660
gtggagcget acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720
atctgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780
tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3840
cacagcetga tegaggagag ecagaaceag caggagaaga acgagcagga getgetggag 3900
ctggacaagt gggccagcct gtggaactgg ttcaacatca ccaactggct gtggtacatc 3960
aagctgttca tcatgatcgt gggcggcctg gtgggcctgc gcatcgtgtt cgccgtgctg 4020
agcatctgag gatccagatc tgctgtgcct tctagttgcc agccatctgt tgtttgcccc 4080
tececegtge etteettgae eetggaaggt gecaetecea etgteettte etaataaaat 4140
caggacagca agggggagga ttggggaagac aatagcaggc atgctgggga tgcggtgggc 4260
tctatgggta cccaggtgct gaagaattga cccggttcct cctgggccag aaagaagcag 4320
gcacatecee ttetetgtga cacaceetgt ecaegeeeet ggttettagt tecageeeca 4380
ctcataggac actcatagct caggagggct ccgccttcaa tcccacccgc taaagtactt 4440
ggagcggtet etecetecet cateageeca ecaaaccaaa eetageetee aagagtggga 4500
agaaattaaa gcaagatagg ctattaagtg cagagggaga gaaaatgcct ccaacatgtg 4560
aggaagtaat gagagaaatc atagaatttt aaggccatca tggccttaat cttccgcttc 4620
ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagctcactc 4680
aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 4740
aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 4800
gctccgccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc 4860
gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt 4920
tecgaceetg eegettaceg gatacetgte egeetttete eettegggaa gegtggeget 4980
ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg 5040
ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct 5100
tgagtccaac ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat 5160
tagcagageg aggtatgtag geggtgetac agagttettg aagtggtgge ctaactaegg 5220
ctacactaga agaacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa 5280
aagagttggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gtttttttgt 5340
ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 5400
```

```
tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt 5460
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 5520
aagtatatat gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat 5580
ctcagcgatc tgtctatttc gttcatccat agttgcctga ctcggggggg ggggcgctq 5640
aggtctgcct cgtgaagaag gtgttgctga ctcataccag gcctgaatcg ccccatcatc 5700
cagccagaaa gtgagggagc cacggttgat gagagctttg ttgtaggtgg accagttgqt 5760
gattltgaac tittgctttg ccacggaacg gtctqcqttq tcqqqaaqat qcqtqatctq 5820
atcettcaac tcagcaaaag ttegatttat tcaacaaage egeegteeeg tcaagtcage 5880
gtaatqctct gccagtgtta caaccaatta accaattctg attagaaaaa ctcatcgagc 5940
atcaaatgaa actgcaattt attcatatca ggattatcaa taccatattt ttgaaaaagc 6000
cgtttctgta atgaaggaga aaactcaccg aggcagttcc ataggatggc aagatcctgg 6060
tateggtetg egatteegae tegteeaaca teaatacaac etattaattt eecetegtea 6120
aaaataaggt tatcaagtga gaaatcacca tgagtgacga ctgaatccgg tgagaatggc 6180
aaaagettat geatttettt ceagaettgt teaacaggee agecattaeg etegteatea 6240
aaatcactcg catcaaccaa accgttattc attogtgatt gcgcctgagc gagacgaaat 6300
acgcgatcgc tgttaaaagg acaattacaa acaggaatcg aatgcaaccg gcgcaggaac 6360
actgccageg catcaacaat attttcacct gaatcaggat attcttctaa tacctggaat 6420
gctgtttttcc cggggatcgc agtggtgagt aaccatgcat catcaggagt acggataaaa 6480
gtaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctqq cqcatcqqqc 6600
ttcccataca atcgatagat tgtcgcacct gattgcccga cattatcgcg agcccattta 6660
tacccatata aatcagcatc catgttggaa tttaatcgcg gcctcqagca agacgtttcc 6720
cgttgaatat ggctcataac accccttgta ttactgttta tgtaagcaga cagttttatt 6780
gttcatgatg atatattttt atcttgtgca atgtaacatc agagattttg agacacaacg 6840
tggctttccc cccccccca ttattgaagc atttatcagg gttattgtct catgagcgga 6900
tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 6960
aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg 7020
cgtatcacga ggccctttcg to
```

```
<210> 10
<211> 6859
<212> DNA
<213> Artificial Sequence
<223> plasmid pVR1012x/s containing HIV genes
<400> 10
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtq 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaqattgq 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cocgcetggc tgaccgccca acgacccccg cccattgacg tcaataatga cqtatqttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
```

|            |            |            |              |              |                          | 001,20, |
|------------|------------|------------|--------------|--------------|--------------------------|---------|
| agctcgttta | gtgaaccgto | agatcgcctg | gagacgccat   | ccacgctgtt   | : ttgacctcca             | 960     |
| tagaagacac | cgggaccgat | ccagcctccg | cggccgggaa   | cggtgcattg   | gaacgcggat               | 1020    |
| teceegtgee | aagagtgacg | taagtaccgc | ctatagacto   | : tataggcaca | cccctttggc               | 1080    |
| tcttatgcat | gctatactgt | ttttggcttg | gggcctatac   | acccccgctt   | ccttatgcta               | 1140    |
| taggtgatgg | tatagcttag | cctataggtg | ' tgggttattg | accattattg   | accactcccc               | 1200    |
| tattggtgac | gatactttcc | attactaatc | cataacatgg   | ctctttgcca   | caactatctc               | 1260    |
| tattggctat | atgccaatac | tctgtccttc | agagactgac   | acggactctg   | tatttttaca               | 1320    |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca   | acaacgccgt   | ccccgtgcc                | 1380    |
| cgcagttttt | attaaacata | gcgtgggatc | tccacgcgaa   | tctcgggtac   | gtgttccgga               | 1440    |
| catgggctct | tctccggtag | cggcggagct | tccacatccg   | agccctggtc   | ccatgcctcc               | 1500    |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag   | tggaggccag   | acttaggcac               | 1560    |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg   | tggcggtagg   | gtatgtgtct               | 1620    |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag   | atggaagact   | taaggcagcg               | 1680    |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat   | aagagtcaga   | ggtaactccc               | 1740    |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag   | cagtactcgt   | tgctgccgcg               | 1800    |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt   | tcctttccat   | gggtcttttc               | 1860    |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg   | ccgctctaga   | caccatgcgc               | 1920    |
| gtgaaggaga | agtaccagca | cctgtggcgc | tggggctggc   | gctggggcac   | catgctgctg               | 1980    |
| ggcatgctga | tgatctgcag | cgccaccgag | aagctgtggg   | tgaccgtgta   | ctacggcgtg               | 2040    |
| cccgtgtgga | aggaggccac | caccaccctg | ctctgcgcca   | gcgacgccaa   | ggcctacgac               | 2100    |
| accgaggtgc | acaacgtgtg | ggccacccac | gcctgcgtgc   | ccaccgaccc   | caacccccag               | 2160    |
| gaggtggtgc | tggtgaacgt | gaccgagaac | ttcgacatgt   | ggaagaacga   | catggtggag               | 2220    |
| cagatgcacg | aggacatcat | cagcctgtgg | gaccagagcc   | tgaagccctg   | cgtgaagctg               | 2280    |
| acccccctgt | gcgtgagcct | gaagtgcacc | gacctgaaga   | acgacaccaa   | caccaacagc               | 2340    |
| agcagcggcc | gcatgatcat | ggagaagggc | gagatcaaga   | actgcagctt   | caacatcagc               | 2400    |
| accagcatcc | gcggcaaggt | gcagaaggag | tacgccttct   | tctacaagct   | ggacatcatc               | 2460    |
| cccatcgaca | acgacaccac | cagctacagc | ctgaccagct   | gcaacaccag   | cgtgatcacc               | 2520    |
| caggcctgcc | ccaaggtgag | cttcgagccc | atccccaacc   | actactgcgc   | ccccgccggc               | 2580    |
| ttcaccatcc | tgaagtgcaa | caacaagacc | ttcaacggca   | ccggcccctg   | caccaacgtg               | 2640    |
| agcaccgtgc | agtgcaccca | cggcatccgc | cccgtggtga   | gcacccagct   | gctgctgaac               | 2700    |
| ggeageetgg | ccgaggagga | ggtggtgatc | cgcagcgtga   | acttcaccga   | caacgccaag               | 2760    |
| accateateg | rgcagergaa | caccagegeg | gagatcaact   | gcacccgccc   | caacaacaac               | 2820    |
| accogcaage | bassassass | ccagegegge | cccggccgcg   | ccttcgtgac   | categgeaag               | 2880    |
| accygcaaca | cgcgccaggc | ccactgcaac | accageegeg   | ccaagtggaa   | caacaccctg               | 2940    |
| cagaagaaga | acageaaget | gegegageag | ttcggcaaca   | acaagaccat   | catcttcaag               | 3000    |
| ttctactcca | geggegaeee | cgagategeg | acccacaget   | tcaactgcgg   | cggcgagttc               | 3060    |
| gagggagg   | acagcaccca | getgeteaae | agcacctggt   | ccaacagcac   | ctggagcacc               | 3120    |
| atcatcaaca | tataggagga | gggcagcgac | accatcaccc   | tgccctgccg   | catcaagcag               | 3180    |
| atcaccaaca | gcagcagaa  | ggcgggcaag | gecatgracg   | cccccccat    | cagcggccag               | 3240    |
| aacgagaggg | agatetteeg | categgeeeg | aggazastaa   | gegaeggegg   | caacagcaac               | 3300    |
| ctgtacaagt | acaaggtggt | gaagatcgag | ggcgacatgc   | bagagagaga   | gcgcagcgag<br>caaggccaag | 3360    |
| cttaccatac | addcccdcca | gaagacegag | aggategteg   | aggagagaga   | caacctgctg               | 3420    |
| cacaccatca | aggcccacca | geagetgage | cacctcaccc   | tatagagagat  | caacctgetg               | 3480    |
| caggecegea | tectaaceat | gaaacactac | ctgaaggacg   | ageaggeat    | gcagatctgg               | 3540    |
| aaccacacca | cctagataga | atagascac  | gagatcaaca   | agcagccgga   | cctgatccac               | 3600    |
| agcctgatcg | aggagagga  | gaaccagcag | gagaccaaca   | accacaccag   | gctggagctg               | 3000    |
| gacaagtggg | ccagcctata | gaactggttc | aacatcacca   | actorctoto   | gtacatcaag               | 3720    |
| ctattcatca | tgatcgtggg | caacctaata | gacctacaca   | tcatattcac   | cgtgctgagc               | 3/00    |
| atctgaggat | ccagatetor | tatacettet | agttgccage   | catctottot   | ttgcccctcc               | 3000    |
| cccgtgcctt | ccttgaccct | ggaaggtgcc | actoccacto   | teettteeta   | ataaaatgag               | 3960    |
| gaaattgcat | cgcattqtct | gagtaggtgt | cattctattc   | tagagaahaa   | ggtggggcag               | 4020    |
| gacagcaagq | gggaggatta | qqaaqacaat | agcaggcatg   | ctadaastac   | ggtggggcag               | 4020    |
| atgggtaccc | aggtgctgaa | gaattgaccc | ggttcctcct   | qqqccaqaaa   | gaagcaggca               | 4140    |
| catccccttc | tctgtgacac | accetgteca | cgcccctaat   | tettagttee   | agccccactc               | 4200    |
|            |            | -          |              |              |                          | -200    |

```
ataggacact catageteag gagggeteeg cetteaatee caceegetaa agtaettgga 4260
geggtetete ceteceteat cageccacca aaccaaacct ageetecaag agtgggaaga 4320
aattaaagca agataggcta ttaagtgcag agggagagaa aatgcctcca acatgtgagg 4380
aagtaatgag agaaatcata gaattttaag gccatcatgg ccttaatctt ccgcttcctc 4440
gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa 4500
ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 4560
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct 4620
ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 4680
aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc 4740
gaccetgeeg ettaceggat acetgteege ettteteeet tegggaageg tggegettte 4800
tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg 4860
tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga 4920
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag 4980
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtggccta actacggcta 5040
cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag 5100
agttggtagc tettgateeg geaaacaaac cacegetggt ageggtggtt tttttgtttg 5160
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 5220
ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc 5280
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag 5340
tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 5400
agcgatctgt ctatttcgtt catccatagt tgcctgactc ggggggggg ggcgctgagg 5460
tetgeetegt gaagaaggtg ttgetgacte ataccaggee tgaategeee cateatecag 5520
ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc agttggtgat 5580
cttcaactca gcaaaagttc gatttattca acaaagccgc cgtcccgtca agtcagcgta 5700
atgctctgcc agtgttacaa ccaattaacc aattctgatt agaaaaactc atcgagcatc 5760
aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5820
ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag atcctggtat 5080
cggtctgcga ttccgactcg tccaacatca atacaaccta ttaatttccc ctcqtcaaaa 5940
ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga gaatggcaaa 6000
agettatgea tttettteea gaettgttea acaggeeage cattaegete gteateaaaa 6060
teactegeat caaccaaace gttatteatt egtgattgeg cetgagegag acgaaataeg 6120
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact 6180
gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6240
gttttcccgg ggatcgcagt ggtgagtaac catgcatcat caggagtacg gataaaatgc 6300
ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat ctcatctgta 6360
acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc 6420
ccatacaatc gatagattgt cgcacctgat tgcccgacat tatcgcgagc ccatttatac 6480
ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga cgtttcccgt 6540
tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag ttttattgtt 6600
catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6660
ctttccccc cccccatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 6720
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 6780
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6840
atcacgaggc cctttcgtc
                                                                6859
```

```
<210> 11
<211> 6965
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
```

<400> 11 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaeteeee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggte ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg ageeetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgage gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegecacea gacataatag etgacagaet aacagaetgt teettteeat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga acaccatgcg 1920 cgtgaaggag aagtaccagc acctgtggcg ctggggctgg cgctggggca ccatgctgct 1980 gggcatgctg atgatctgca gcgccaccga gaagctgtgg gtgaccgtgt actacggcgt 2040 gecegtgtgg aaggaggeea ceaceaceet getetgegee agegaegeea aggeetaega 2100 caccgaggtg cacaacgtgt gggccaccca cgcctgcgtg cccaccgacc ccaaccccca 2160 ggaggtggtg ctggtgaacg tgaccgagaa cttcgacatg tggaagaacg acatggtgga 2220 gcagatgcac gaggacatca tcagcctgtg ggaccagagc ctgaagccct gcgtgaagct 2280 gacccccctg tgcgtgagcc tgaagtgcac cgacctgaag aacgacacca acaccaacag 2340 cagcagegge egeatgatea tggagaaggg egagateaag aactgeaget teaacateag 2400 caccagcatc cgcggcaagg tgcagaagga gtacgccttc ttctacaagc tggacatcat 2460 ccccatcgac aacgacacca ccagctacag cctgaccagc tgcaacacca gcgtgatcac 2520 ccaggectge cccaaggtga gettegagee catecccaae caetactgeg cccccgccgg 2580 cttcaccatc ctgaagtgca aggacaagaa gttcaacggc aagggcccct gcaccaacgt 2640 gageaccgtg cagtgcaccc acggcatccg ccccgtggtg agcacccagc tgctgctgaa 2700 cgaccgcctg gccgaggagg aggtggtgat ccgcagcgcc aacttcgccg acaacgccaa 2760 ggtgatcatc gtgcagctga acgagagcgt ggagatcaac tgcacccgcc ccaacaacaa 2820 caccegcaag agcatecaca teggeceegg eegegeette tacaccaceg gegagateat 2880 cggcgacatc cgccaggccc actgcaacct gagccgcgcc aagtggaacg acaccctgaa 2940 caagategtg atcaagetge gegageagtt eggeaacaag accategtgt teaageacag 3000 cageggegge gacceegaga tegtgaceea cagetteaac tgeggeggeg agttetteta 3060 ctgcaacagc acccagctgt tcaacagcac ctggttcaac agcacctgga gcaccgaggg 3120 cagcaacaac accgagggca gcgacaccat caccctgccc tgccgcatca agcagatcat 3180

|            |            |            |            |            | . 101/0    | 301/23/ |
|------------|------------|------------|------------|------------|------------|---------|
| caacatgtgg | cagaaggtgg | gcaaggccat | gtacgccccc | cccatcagcg | gccagatccg | 3240    |
| ctgcagcagc | aacatcaccg | gcctgctgct | gacccgcgac | ggcggcaaca | gcaacaacga | 3300    |
| gagcgagatc | ttccgcctgg | gcggcggcga | catgcgcgac | aactggcgca | gcgagctgta | 3360    |
| caagtacaag | gtggtgaaga | tegageceet | gggcgtggcc | cccaccaagg | ccaagcgccg | 3420    |
| cgtggtgcag | cgcgagaagc | gcgccgtggg | catcggcgcc | ctgttcctgg | gcttcctggg | 3480    |
| cgccgccggc | agcaccatgg | gcgccgccag | catgaccctg | accgtgcagg | cccgccagct | 3540    |
| gctgagcggc | atcgtgcagc | agcagaacaa | cctgctgcgc | gccatcgagg | cccagcagca | 3600    |
| cctgctgcag | ctgaccgtgt | ggggcatcaa | gcagctgcag | gcccgcatcc | tggccgtgga | 3660    |
| gcgctacctg | aaggaccagc | agctgctggg | catctggggc | tgcagcggca | agcttatctg | 3720    |
| caccaccgcc | gtgccctgga | acgccagctg | gagcaacaag | agcctggagc | agatetggaa | 3780    |
| ccacaccacc | tggatggagt | gggaccgcga | gatcaacaac | tacaccagcc | tgatccacag | 3840    |
| cctgatcgag | gagagccaga | accagcagga | gaagaacgag | caggagetge | tggagctgga | 3900    |
| caagtgggcc | agcctgtgga | actggttcaa | catcaccaac | tggctgtggt | gaggatccag | 3960    |
| atctgctgtg | ccttctagtt | gccagccatc | tgttgtttgc | ccctcccccg | tgccttcctt | 4020    |
| gaccctggaa | ggtgccactc | ccactgtcct | ttcctaataa | aatgaggaaa | ttgcatcgca | 4080    |
| ttgtctgagt | aggtgtcatt | ctattctggg | gggtggggtg | gggcaggaca | gcaaggggga | 4140    |
|            |            |            | ggatgcggtg |            |            |         |
| gctgaagaat | tgacccggtt | cctcctgggc | cagaaagaag | caggcacatc | cccttctctg | 4260    |
| tgacacaccc | tgtccacgcc | cctggttctt | agttccagcc | ccactcatag | gacactcata | 4320    |
| gctcaggagg | gctccgcctt | caatcccacc | cgctaaagta | cttggagcgg | tctctccctc | 4380    |
| cctcatcagc | ccaccaaacc | aaacctagcc | tccaagagtg | ggaagaaatt | aaaqcaaqat | 4440    |
| aggctattaa | gtgcagaggg | agagaaaatg | cctccaacat | gtgaggaagt | aatgagagaa | 4500    |
| atcatagaat | tttaaggcca | tcatggcctt | aatcttccgc | ttcctcgctc | actgactcgc | 4560    |
| tgcgctcggt | cgttcggctg | cggcgagcgg | tatcagctca | ctcaaaggcg | gtaatacggt | 4620    |
| tatccacaga | atcaggggat | aacgcaggaa | agaacatgtg | agcaaaaqqc | caqcaaaaqq | 4680    |
| ccaggaaccg | taaaaaggcc | gcgttgctgg | cgtttttcca | taggeteege | cccctqacq  | 4740    |
| agcatcacaa | aaatcgacgc | tcaagtcaga | ggtggcgaaa | cccgacagga | ctataaagat | 4800    |
| accaggcgtt | tccccctgga | agctccctcg | tgcgctctcc | tgttccgacc | ctgccgctta | 4860    |
| ccggatacct | gtccgccttt | ctcccttcgg | gaagcgtggc | gctttctcat | ageteacget | 4920    |
| gtaggtatct | cagttcggtg | taggtcgttc | gctccaagct | gggctgtgtg | cacqaacccc | 4980    |
| ccgttcagcc | cgaccgctgc | gccttatccg | gtaactatcg | tcttgagtcc | aacccqqtaa | 5040    |
| gacacgactt | atcgccactg | gcagcagcca | ctggtaacag | gattagcaga | gcgaggtatg | 5100    |
| taggcggtgc | tacagagttc | ttgaagtggt | ggcctaacta | cggctacact | agaagaacag | 5160    |
| tatttggtat | ctgcgctctg | ctgaagccag | ttaccttcgg | aaaaagagtt | ggtagctctt | 5220    |
| gatccggcaa | acaaaccacc | gctggtagcg | gtggttttt  | tgtttgcaag | cagcagatta | 5280    |
| cgcgcagaaa | aaaaggatct | caagaagatc | ctttgatctt | ttctacgggg | tctgacgctc | 5340    |
| agtggaacga | aaactcacgt | taagggattt | tggtcatgag | attatcaaaa | aggatcttca | 5400    |
| cctagatcct | tttaaattaa | aaatgaagtt | ttaaatcaat | ctaaagtata | tatgagtaaa | 5460    |
| cttggtctga | cagttaccaa | tgcttaatca | gtgaggcacc | tatctcagcg | atctgtctat | 5520    |
| ttcgttcatc | catagttgcc | tgactcgggg | gggggggcg  | ctgaggtctg | cctcgtgaag | 5580    |
| aaggtgttgc | tgactcatac | caggcctgaa | tcgccccatc | atccagccag | aaagtgaggg | 5640    |
| agccacggtt | gatgagagct | ttgttgtagg | tggaccagtt | ggtgattttg | aacttttgct | 5700    |
| ttgccacgga | acggtctgcg | ttgtcgggaa | gatgcgtgat | ctgatccttc | aactcagcaa | 5760    |
| aagttcgatt | tattcaacaa | agccgccgtc | ccgtcaagtc | agcgtaatgc | tctgccagtg | 5820    |
| ttacaaccaa | ttaaccaatt | ctgattagaa | aaactcatcg | agcatcaaat | gaaactgcaa | 5880    |
| tttattcata | tcaggattat | caataccata | tttttgaaaa | agccgtttct | gtaatgaagg | 5940    |
| agaaaactca | ccgaggcagt | tccataggat | ggcaagatcc | tggtatcggt | ctgcgattcc | 6000    |
| gactcgtcca | acatcaatac | aacctattaa | tttcccctcg | tcaaaaataa | ggttatcaag | 6060    |
| tgagaaatca | ccatgagtga | cgactgaatc | cggtgagaat | ggcaaaagct | tatgcatttc | 6120    |
| tttccagact | tgttcaacag | gccagccatt | acgctcgtca | tcaaaatcac | tcgcatcaac | 6180    |
| caaaccgtta | ttcattcgtg | attgcgcctg | agcgagacga | aatacgcgat | cgctgttaaa | 6240    |
| aggacaatta | caaacaggaa | tcgaatgcaa | ccggcgcagg | aacactgcca | gcgcatcaac | 6300    |
| aatattttca | cctgaatcag | gatattcttc | taatacctgg | aatgctgttt | tcccggggat | 6360    |
| cgcagtggtg | agtaaccatg | catcatcagg | agtacggata | aaatgcttga | tggtcggaag | 6420    |
| aggcataaat | tccgtcagcc | agtttagtct | gaccatctca | tctgtaacat | cattggcaac | 6480    |
|            |            |            |            |            |            |         |

gctacetttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6540

```
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6600
atccatgttg gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6660
aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt 6720
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6780
ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6840
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6900
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6960
tcgtc
<210> 12
<211> 6784
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 12
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaceteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eccetttgge 1080
tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accaptattg accaptcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
egegecacea gacataatag etgacagaet aacagaetgt teettteeat gggtetttte 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
głgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
```

cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220 cagatgcacg aggacatcat cagcotgtgg gaccagagoo tgaagoootg cgtgaagotg 2280 acceccetgt gegtgageet gaagtgeace gacetgaaga acgaeaceaa caccaacage 2340 agcageggee geatgateat ggagaaggge gagateaaga actgeagett caacateage 2400 accagcatec geggeaaggt geagaaggag taegeettet tetacaaget ggacateate 2460 cccategaca acgacaceae cagetacage etgaceaget gcaacaceag egtgateace 2520 caggootgoo ccaaggtgag ottogagooc atocccaaco actactgogo coccgeoggo 2580 ttegecatee tgaagtgeaa ggacaagaag tteaaeggea agggeeeetg caccaaegtg 2640 ageacegtge agtgeaceca eggeateege eeegtggtga geacecaget getggttaeg 2700 ggtaacctgg ccgaggagga ggtggtgatc cgcagcgcta acttcgccga caacgccaaq 2760 gtgatcatcg tgcagctgaa cgagagcgtg gagatcaact gcacccgccc caacaacaac 2820 accegeaaga geatecacat eggeecegge egegeettet acaecacegg egagateate 2880 ggcgacatcc gccaggccca ctgcaacctg agccgcgcca agtggaacga caccctgaac 2940 aagatcgtga tcaagctgcg cgagcagttc ggcaacaaga ccatcgtgtt caagcacagc 3000 ageggeggeg acceegagat egtgaceeae agetteaaet geggeggega gttettetae 3060 tgcaacagca cccagctgtt caacagcacc tggttcaaca gcacctggag caccgagggc 3120 agcaacaaca ccgagggcag cgacaccatc accetgeeet geegcateaa gcagateate 3180 aacatgtggc agaaggtggg caaggccatg tacgccccc ccatcagcgg ccaqatccqc 3240 tgcagcagca acatcaccgg cetgetgetg accegegacg geggeaacag caacaacgaq 3300 agegagatet teegeetggg eggeggegae atgegegaea aetggegeag egagetgtae 3360 aagtacaagg tggtgaagat cgagcccctg ggcgtggccc ccaccaaggc caagcttacc 3420 gtccaggccc gccagctgct gagcggcatc gtgcagcagc agaacaacct gctgcgcgcc 3480 ategaggeee ageageacet getgeagetg acegtgtggg geateaagea getgeaggee 3540 cgcaccetgg ccgtggagcg ctacctgaag gaccagcagc tgctcgagca gatctggaac 3600 cacaccacct ggatggagtg ggaccgcgag atcaacaact acaccagcct gatccacagc 3660 ctgatcgagg agagccagaa ccagcacgag aagaacgagc aggagctgct ggagctggac 3720 aagtgggcca gcctgtggaa ctggttcaac atcaccaact ggctgtggtg aggatccaga 3780 tetgetgtge ettetagttg ceagecatet gttgtttgce cetececegt geetteettg 3840 accetggaag gtgccactee caetgteett teetaataaa atgaggaaat tgcategeat 3900 gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatggg tacccaggtg 4020 ctgaagaatt gacccggttc ctcctgggcc agaaagaagc aggcacatcc ccttctctgt 4080 gacacaccct gtccacgccc ctggttctta gttccagccc cactcatagg acactcatag 4140 ctcaggaggg ctccgccttc aatcccaccc gctaaagtac ttggagcggt ctctccctcc 4200 ctcatcagcc caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata 4260 ggctattaag tgcagaggga gagaaaatgc ctccaacatg tgaggaagta atgagagaaa 4320 teatagaatt ttaaggeeat eatggeetta atetteeget teetegetea etgaeteget 4380 gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt 4440 atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 4500 caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 4560 gcatcacaaa aatcgacget caagtcagag gtggcgaaac ccgacaggac tataaagata 4620 ccaggogtht ccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 4680 eggatacetg teegeettte teeetteggg aagegtggeg ettteteata geteaegetg 4740 taggtatete agtteggtgt aggtegtteg etecaagetg ggetgtgtge aegaaceee 4800 cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 4860 acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 4920 aggeggtget acagagttet tgaagtggtg geetaactae ggetacaeta gaagaacagt 4980 atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 5040 atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 5100 gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacqqqqt ctqacqctca 5160 gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 5220 ctagateett ttaaattaaa aatgaagttt taaateaate taaagtatat atgagtaaae 5280 ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 5340

```
tegtteatee atagttgeet gaeteggggg gggggggege tgaggtetge etegtgaaga 5400
aggtgttgct gactcatacc aggcctgaat cgccccatca tccagccaga aagtgaggga 5460
gccacggttg atgagagctt tgttgtaggt ggaccagttg gtgattttga acttttgctt 5520
tgccacggaa cggtctgcgt tgtcgggaag atgcgtgatc tgatccttca actcagcaaa 5580
agttcgattt attcaacaaa gccgccgtcc cgtcaagtca gcgtaatgct ctgccagtgt 5640
tacaaccaat taaccaatte tgattagaaa aactcatega geatcaaatg aaactgeaat 5700
ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg taatgaagga 5760
gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg 5820
actogtocaa catcaataca acctattaat ttoccotogt caaaaataag gttatcaagt 5880
gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt atgcatttct 5940
ttccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact cgcatcaacc 6000
aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgatc gctgttaaaa 6060
ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag cgcatcaaca 6120
atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt cccggggatc 6180
gcagtggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat ggtcggaaga 6240
ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc attggcaacg 6300
ctacctttgc catgtttcag aaacaactct ggcgcatcgg gcttcccata caatcgatag 6360
attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata taaatcagca 6420
tecatgttgg aatttaateg eggeetegag caagaegttt eeegttgaat atggeteata 6480
acaccccttg tattactgtt tatgtaagca gacagtttta ttgttcatga tgatatattt 6540
ttatettgtg caatgtaaca teagagattt tgagacacaa egtggettte ecceecece 6600
cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 6660
tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 6720
taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 6780
cqtc
                                                                  6784
```

```
<210> 13
<211> 7037
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 13
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
```

cgtcaatggg agtttgttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 tccccgtgcc aagatgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080

```
tettatgeat getatactgt ttttggettg gggeetatae acceegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg ageeetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga acaccatgcg 1920
cgtgaaggag aagtaccagc acctgtggcg ctggggctgg cgctggggca ccatgctgct 1980
gggcatgctg atgatctgca gcgccaccga gaagctgtgg gtgaccgtgt actacggcgt 2040
gcccgtgtgg aaggaggcca ccaccacct gctctgcgcc agcgacgcca aggcctacga 2100
caccgaggtg cacaacgtgt gggccaccca cgcctgcgtg cccaccgacc ccaacccca 2160
ggaggtggtg ctggtgaacg tgaccgagaa cttcgacatg tggaagaacg acatggtgga 2220
gcagatgcac gaggacatca tcagcctgtg ggaccagagc ctgaagccct gcgtgaagct 2280
gaccccctg tgcgtgagcc tgaagtgcac cgacctgaag aacgacacca acaccaacag 2340
cagcagegge egeatgatea tggagaaggg egagateaag aactgeaget teaacateag 2400
caccagcate egeggeaagg tgeagaagga gtacgeette ttetacaage tggacateat 2460
ccccatcgac aacgacacca ccagctacag cctgaccagc tgcaacacca gcgtgatcac 2520
ccaggcctgc cccaaggtga gcttcgagcc catccccaac cactactgcg cccccgccgg 2580
cttcaccatc ctgaagtgca aggacaagaa gttcaacggc aagggcccct gcaccaacgt 2640
gagcaccgtg cagtgcaccc acggcatccg ccccgtggtg agcacccagc tgctgctgaa 2700
cgaccgcctg gccgaggagg aggtggtgat ccgcagcgcc aacttcgccg acaacgccaa 2760
ggtgatcatc gtgcagctga acgagagcgt ggagatcaac tgcacccgcc ccaacaacaa 2820
caccegcaag agcatecaca teggeceegg eegegeette tacaceaeeg gegagateat 2880
cggcgacatc cgccaggccc actgcaacct gagccgcgcc aagtggaacg acaccctgaa 2940
caagatcgtg atcaagctgc gcgagcagtt cggcaacaag accatcgtgt tcaagcacag 3000
cageggegge gacceegaga tegtgaceea cagetteaac tgeggeggeg agttetteta 3060
ctgcaacagc acccagctgt tcaacagcac ctggttcaac agcacctgga gcaccgaggg 3120
cagcaacaac accgagggca gcgacaccat caccctgccc tgccgcatca agcagatcat 3180
caacatgtgg cagaaggtgg gcaaggccat gtacgcccc cccatcagcg gccagatccg 3240
ctgcagcagc aacatcaccg gcctgctgct gacccgcgac ggcggcaaca gcaacaacga 3300
gagcgagate ttccgcctgg gcggcggcga catgcgcgac aactggcgca gcgagctgta 3360
caagtacaag gtggtgaaga tcgagcccct gggcgtggcc cccaccaagg ccaagcgccg 3420
cgtggtgcag cgcgagaagc gcgccgtggg catcggcgcc ctgttcctgg gcttcctggg 3480
egeegeegge ageaceatgg gegeegeeag catgaceetg accgtgeagg ceegeeaget 3540
gctgagcggc atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca 3600
cctgctgcag ctgaccgtgt ggggcatcaa gcagctgcag gcccgcatcc tggccgtgga 3660
gegetacetg aaggaceage agetgetggg catctgggge tgeageggea agettatetg 3720
caccaccgcc gtgccctgga acgccagctg gagcaacaag agcctggagc agatctggaa 3780
ccacaccacc tggatggagt gggaccgcga gatcaacaac tacaccagcc tgatccacag 3840
cctgatcgag gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga 3900
caagtgggcc agcctgtgga actggttcaa catcaccaac tggctgtggt acatcaagct 3960
gttcatcatg atcgtgggcg gcctggtggg cctgcgcatc gtgttcgccg tgctgagcat 4020
ctgaggatcc agatetgetg tgcettetag ttgccageca tetgttgttt geceeteeec 4080
cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ctttcctaat aaaatgagga 4140
aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga 4200
cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat 4260
gggtacccag gtgctgaaga attgacccgg ttcctcctgg gccagaaaga agcaggcaca 4320
teceettete tgtgacacae cetgtecaeg eccetggtte ttagttecag ecceaeteat 4380
```

```
aggacactca tagctcagga gggctccgcc ttcaatccca cccgctaaag tacttggagc 4440
ggtctctccc tccctcatca gcccaccaaa ccaaacctag cctccaagag tgggaagaaa 4500
ttaaagcaag ataggctatt aagtgcagag ggagagaaaa tgcctccaac atgtgaggaa 4560
gtaatgagag aaatcataga attttaaggc catcatggcc ttaatcttcc gcttcctcgc 4620
teactgacte getgegeteg gtegttegge tgeggegage ggtateaget caeteaaagg 4680
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 4740
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 4800
gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 4860
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 4920
ecctgceget taceggatae etgteegeet ttetecette gggaagegtg gegettete 4980
atageteaeg etgtaggtat eteagttegg tgtaggtegt tegeteeaag etgggetgtg 5040
tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 5100
ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 5160
gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 5220
ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 5280
ttggtagete ttgateegge aaacaaacea eegetggtag eggtggtttt tttgtttgca 5340
agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 5400
ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 5460
aaaggatett cacetagate ettttaaatt aaaaatgaag ttttaaatea atetaaagta 5520
tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag 5580
cgatctgtct atttcgttca tccatagttg cctgactcgg ggggggggg cgctgaggtc 5640
tgcctcgtga agaaggtgtt gctgactcat accaggcctg aatcgcccca tcatccagcc 5700
agaaagtgag ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt 5760
tgaacttttg ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatcct 5820
teaacteage aaaagttega tttatteaac aaageegeeg teeegteaag teagegtaat 5880
gctctgccag tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa 5940
atgaaactgc aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt 6000
ctgtaatgaa ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg 6060
gtctgcgatt ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat 6120
aaggttatca agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag 6180
cttatgcatt tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc 6240
actegeatea accaaacegt tatteatteg tgattgegee tgagegagae gaaataegeg 6300
ategetgtta aaaggacaat tacaaacagg aategaatge aaceggegea ggaacaetge 6360
cagcgcatca acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt 6420
tttcccgggg atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaatgctt 6480
gatggtcgga agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac 6540
atcattggca acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc 6600
atacaatcga tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc 6660
atataaatca gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg 6720
aatatggctc ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca 6780
tgatgatata tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct 6840
ttccccccc ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat 6900
atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 6960
gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat 7020
cacgaggccc tttcgtc
```

```
<210> 14
<211> 6856
<212> DNA
<213> Artificial Sequence
<220>
```

<223> plasmid pVR1012x/s containing HIV genes

<400> 14 tegegegitt eggigatgae ggitgaaaace tetgacacat geageteeeg gagaeggiea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatabgga gttccgcgtt acataactta cqqtaaatqq 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacqqtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 . catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattqa 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tggaqqtct atataaqcaq 900 agetegitta gigaacegic agategeeig gagaegeeat ceaegeigtt tigaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecce 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae aeggactetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ecatgeetee 1500 aggggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatqcqc 1920 gtgaaggaga agtaccagca cetgtggege tggggetgge getggggeae catgetgetg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040 cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280 acceccetgt gegtgageet gaagtgeace gacetgaaga acgaeaceaa caceaacage 2340 agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400 accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460 cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cqtqatcacc 2520 caggoetgee ceaaggtgag ettegageee atececaace actactgege eccegeegge 2580 ttcgccatcc tgaagtgcaa ggacaagaag ttcaacggca agggcccctg caccaacgtg 2640 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctggttacg 2700 ggtaacctgg ccgaggagga ggtggtgatc cgcagcgcta acttcgccga caacqccaaq 2760 gtgatcatcg tgcagctgaa cgagagcgtg gagatcaact gcacccgccc caacaacaac 2820 accegeaaga geatecacat eggeeeegge egegeettet acaceacegg egagateate 2880 ggcgacatcc gccaggccca ctgcaacctg agccgcgcca agtggaacga caccctqaac 2940 aagategtga teaagetgeg egageagtte ggeaacaaga eeategtgtt caageacage 3000 ageggeggeg acceegagat egtgacceae agetteaact geggeggega gttettetae 3060 tgcaacagca cccagctgtt caacagcacc tggttcaaca gcacctggag caccgaggc 3120

agcaacaaca ccgagggcag cgacaccatc accctgccct gccgcatcaa gcagatcatc 3180

```
aacatgtggc agaaggtggg caaggccatg tacgccccc ccatcagcgg ccagatccgc 3240
tgcagcagca acatcaccgg cctgctgctg acccgcgacg gcggcaacag caacaacgag 3300
agcgagatet teegeetggg eggeggegae atgegegaea aetggegeag egagetgtae 3360
aagtacaagg tggtgaagat cgagcccetg ggcgtggccc ccaccaaggc caagcttacc 3420
gtccaggccc gccagctgct gagcggcatc gtgcagcagc agaacaacct gctgcgcgcc 3480
atcgaggece ageageacet getgeagetg accgtgtggg geatcaagea getgeaggee 3540
cgcaccctgg ccgtggagcg ctacctgaag gaccagcagc tgctcgagca gatctggaac 3600
cacaccacct ggatggagtg ggaccgcgag atcaacaact acaccagcct gatccacagc 3660
ctgatcgagg agagccagaa ccagcacgag aagaacgagc aggagctgct ggagctggac 3720
aagtgggcca gcctgtggaa ctggttcaac atcaccaact ggctgtggta catcaagctg 3780
ttcatcatga tcgtgggcgg cctggtgggc ctgcgcatcg tgttcgccgt gctgagcatc 3840
tgaggateca gatetgetgt geettetagt tgecagecat etgttgtttg eeeeteeee 3900
gtgccttcct tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa 3960
agcaaggggg aggattggga agacaatagc aggcatgctg gggatgcggt gggctctatg 4080
ggtacccagg tgctgaagaa ttgacccggt tcctcctggg ccagaaagaa gcaggcacat 4140
ccccttctct gtgacacacc ctgtccacgc ccctggttct tagttccagc cccactcata 4200
ggacactcat agetcaggag ggetcegeet teaatcecae eegetaaagt aettggageg 4260
gtototocot coctoatoag cocaccaaac caaacctago otocaagagt gggaagaaat 4320
taaagcaaga taggctatta agtgcagagg gagagaaaat gcctccaaca tgtgaggaag 4380
taatgagaga aatcatagaa ttttaaggcc atcatggcct taatcttccg cttcctcgct 4440
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 4500
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 4560
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 4620
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacaqg 4680
actataaaga taccaggcgt ttecceetgg aageteeete gtgegetete etgtteegae 4740
cetgeegett aceggatace tgteegeett tetecetteg ggaagegtgg egetttetea 4800
tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 4860
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 4920
caacceggta agacacgact tategecact ggcagcagec actggtaaca ggattagcag 4980
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 5040
tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 5100
tggtagetet tgateeggea aacaaaceae egetggtage ggtggttttt ttgtttgcaa 5160
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 5220
gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 5280
aaggatette acctagatee tittaaatta aaaatgaagt titaaateaa tetaaagtat 5340
atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 5400
gatetgteta tttegtteat ceatagttge etgaeteggg ggggggggg getgaggtet 5460
gcctcgtgaa gaaggtgttg ctgactcata ccaggcctga atcgccccat catccagcca 5520
gaaagtgagg gagccacggt tgatgagagc tttgttgtag gtggaccagt tggtgatttt 5580
gaacttttgc tttgccacgg aacggtctgc gttgtcggga agatgcgtga tctgatcctt 5640
caactcagca aaagttcgat ttattcaaca aagccgccgt cccgtcaagt cagcgtaatg 5700
ctctgccagt gttacaacca attaaccaat tctgattaga aaaactcatc gagcatcaaa 5760
tgaaactgca atttattcat atcaggatta tcaataccat atttttgaaa aagccgtttc 5820
tgtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg 5880
tetgegatte egactegtee aacateaata caacetatta attteeete gteaaaaata 5940
aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc 6000
ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca 6060
ctcgcatcaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg aaatacgcga 6120
tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgcag gaacactgcc 6180
agcgcatcaa caatattttc acctgaatca ggatattctt ctaatacctg gaatgctgtt 6240
ttcccgggga tcgcagtggt gagtaaccat gcatcatcag gagtacggat aaaatgcttg 6300
atggtcggaa gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca 6360
tcattggcaa cgctaccttt gccatgtttc agaaacaact ctggcgcatc gggcttccca 6420
tacaatcgat agattgtcgc acctgattgc ccgacattat cgcgagccca tttataccca 6480
```

```
tataaatcag catccatgtt ggaatttaat cgcggcctcg agcaagacgt ttcccgttga 6540
atatggctca taacacccct tgtattactg tttatgtaag cagacagttt tattgttcat 6600
gatgatatat ttttatcttg tgcaatgtaa catcagagat tttgagacac aacgtggctt 6660
teceeecce eccattattg aagcatttat cagggttatt gteteatgag eggatacata 6720
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 6780
ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc 6840
acgaggeeet ttegte
<210> 15
<211> 6505
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegttt eggtgatgae ggtgaaaace tetgacaeat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac egggacegat ecageeteeg eggeegggaa eggtgeattg gaacgeggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080
tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaeteece 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag gctagcagcc tgaagtgcac cgacctgaag 2040
aacgacacca acaccaacag cagcagege egeatgatea tggagaaggg cgagateaag 2100
aactgcagct tcaacatcag caccagcatc cgcggcaagg tgcagaagga gtacgccttc 2160
```

| <b>**</b> - * |              |              |              |              |              |      |
|---------------|--------------|--------------|--------------|--------------|--------------|------|
| rccracaage    | c tggacatcat | : ccccatcgad | c aacgacacca | a ccagctaca  | g cctgaccago | 2220 |
| cgcaacacca    | a gcgtgatcac | : ccaggcctgo | cccaaggtg:   | a gettegage  | catececaae   | 2280 |
| cactactgc     | g cccccgccgg | , cttcgccatc | : ctgaagtgca | a aggacaagaa | a gttcaacgg  | 2340 |
| aagggcccci    | gcaccaacgt   | gagcaccgtg   | g cagtgcacco | acggcatccg   | g ccccataato | 2400 |
| agcacccago    | : tgctgctgaa | ı cggtagcctg | gccgaggagg   | g aggtggtgat | cogcagoget   | 2460 |
| aacttcgcc     | g acaacgccaa | ı ggtgatcato | : gtgcagctga | a acgagagcgt | ggagatcaac   | 2520 |
| rgcacccgc     | c ccaacaacaa | cacccgcaag   | , agcatccaca | a teggeecegg | g ccgcgccttc | 2580 |
| cacaccacco    | g gcgagatcat | . cggcgacatc | cgccaggccc   | c actgcaacct | gageegege    | 2640 |
| aagtggaacg    | g acaccctgaa | caagatcgtg   | atcaagetge   | gcgagcagtt   | cggcaacaag   | 2700 |
| accatcgtgt    | : tcaagcacag | cageggegge   | gaccccgaga   | togtgaccca   | cagetteaac   | 2760 |
| tgcggcggcg    | , agttcttcta | ctgcaacago   | acccagetgt   | tcaacagcac   | ctggttcaac   | 2820 |
| agcacctgga    | . gcaccgaggg | cagcaacaac   | accgagggca   | gcgacaccat   | caccetgeee   | 2880 |
| tgccgcatca    | agcagatcat   | caacatgtgg   | cagaaggtgc   | gcaaggccat   | gtacgcccc    | 2940 |
| cccatcagcg    | gccagatccg   | ctgcagcagc   | aacatcacco   | gcctactact   | gacccgcgac   | 3000 |
| ggcggcaaca    | gcaacaacga   | gagcgagatc   | ttccqtccqq   | gcaacaacaa   | catgcgcgac   | 3060 |
| aactggcgca    | gcgagctgta   | caagtacaag   | gtggtgaaga   | tcgagecect   | gggcgtggcc   | 3120 |
| cccaccaagg    | ccaagettae   | cgtccaggcc   | cgccaqctqc   | tgagcggcat   | cgtgcagcag   | 3180 |
| cagaacaaco    | tgctgcgcgc   | catcgaggcc   | cagcagcacc   | tactacaact   | gaccgtgtgg   | 3240 |
| ggcatcaago    | agctgcaggc   | ccgcaccctg   | gccgtggagc   | gctacctgaa   | ggaccagcag   | 3300 |
| cagetegage    | agatctggaa   | ccacaccacc   | tagatagagt   | gggaccgcga   | gatcaacaac   | 3360 |
| tacaccagco    | tgatccacag   | cctgatcgag   | gagagccaga   | accagcagga   | gaagaacgag   | 3420 |
| caggagetge    | tggagctgga   | caagtgggcc   | agcctgtgga   | actoottcaa   | catcaccaac   | 3420 |
| tggctgtggt    | gaggatccag   | atctgctgtg   | ccttctagtt   | gccagccatc   | tgttgtttgc   | 3540 |
| ccctcccccg    | tgccttcctt   | gaccctggaa   | ggtgccactc   | ccactatect   | ttcctaataa   | 3600 |
| aatgaggaaa    | ttgcatcgca   | ttgtctgagt   | aggtgtcatt   | ctattctggg   | gggtggggtg   | 3660 |
| gggcaggaca    | gcaaggggga   | ggattgggaa   | gacaatagca   | gacatactag   | ggatgcggtg   | 3720 |
| ggctctatgg    | gtacccaggt   | gctgaagaat   | tgacccggtt   | cctcctaaac   | cagaaagaag   | 3790 |
| caggcacatc    | cccttctctg   | tgacacaccc   | tatccacacc   | cctaattett   | agttccagcc   | 3840 |
| ccactcatag    | gacactcata   | gctcaggagg   | gctccgcctt   | caatcccacc   | cgctaaagta   | 3900 |
| cttggagcgg    | tctctccctc   | cctcatcage   | ccaccaaacc   | aaacctagcc   | tccaagagtg   | 3960 |
| ggaagaaatt    | aaagcaagat   | aggctattaa   | gtgcagaggg   | agagaaaatg   | cctccaacat   | 4020 |
| gtgaggaagt    | aatgagagaa   | atcatagaat   | tttaaggcca   | tcatggcctt   | aatcttccgc   | 4020 |
| ttcctcgctc    | actgactcgc   | tgcgctcgqt   | cattcaacta   | Caacasacaa   | tatcagctca   | 4140 |
| ctcaaaggcg    | gtaatacggt   | tatccacaga   | atcaggggat   | aacgcaggaa   | agaacatgtg   | 4200 |
| agcaaaaggc    | cagcaaaagg   | ccaggaaccg   | taaaaaqqcc   | acattactaa   | cgtttttcca   | 4260 |
| taggctccgc    | cccctgacg    | agcatcacaa   | aaatcgacgc   | tcaagtcaga   | ggtggcgaaa   | 4320 |
| cccgacagga    | ctataaagat   | accaggcgtt   | tececetaga   | agctccctcg   | tgcgctctcc   | 4380 |
| tgttccgacc    | ctgccgctta   | ccggatacct   | gtccgccttt   | ctcccttcaa   | gaagcgtggc   | 4440 |
| gctttctcat    | agctcacgct   | gtaggtatct   | cagttcggtq   | taggtcgttc   | gctccaagct   | 4500 |
| gggctgtgtg    | cacgaacccc   | ccgttcagcc   | cgaccqctqc   | gccttatccg   | gtaactatcg   | 4560 |
| tcttgagtcc    | aacccggtaa   | gacacgactt   | atcgccactg   | gcagcagcca   | ctggtaacag   | 4620 |
| gattagcaga    | gcgaggtatg   | taggcggtgc   | tacagagttc   | ttgaagtggt   | ggcctaacta   | 4680 |
| cggctacact    | agaagaacag   | tatttggtat   | ctgcgctctq   | ctgaagccag   | ttaccttcoo   | 4740 |
| aaaaagagtt    | ggtagctctt   | gateeggeaa   | acaaaccacc   | gctggtagcg   | ataattttt    | 4800 |
| ţgtttgcaag    | cagcagatta   | cgcgcagaaa   | aaaaqqatct   | caagaagatc   | ctttgatctt   | 4860 |
| ttctacgggg    | tctgacgctc   | agtggaacga   | aaactcacqt   | taagggattt   | tggtcatgag   | 4920 |
| attatcaaaa    | aggatettea   | cctagatcct   | tttaaattaa   | aaatgaagtt   | ttaaatcaat   | 4980 |
| ctaaagtata    | tatgagtaaa   | cttggtctga   | cagttaccaa   | tacttaatca   | gtgaggcacc   | 5040 |
| tatctcagcg    | atctgtctat   | ttcgttcatc   | catagttqcc   | tgactcggg    | gggggggggg   | 5100 |
| ctgaggtctg    | cctcgtgaag   | aaggtgttgc   | tgactcatac   | Caddccpdaa   | tcgccccatc   | 5160 |
| atccagccag    | aaagtgaggg   | agccacqqtt   | gatgagaget   | ttattataaa   | tggaccagtt   | 5220 |
| ggtgattteg    | aacttttgct   | ttgccacgga   | acggtctgcg   | ttatcaaaaa   | gatgcgtgat   | 5280 |
| cigatectte    | aactcagcaa   | aagttcgatt   | tattcaacaa   | agccgccatc   | ccatcaaatc   | 5340 |
| agegeaatge    | tetgecagtg   | ttacaaccaa   | ttaaccaatt   | Ctgattagaa   | aaactcatco   | 5400 |
| agcatcaaat    | gaaactgcaa   | tttattcata   | tcaggattat   | Caataccata   | tttttgaaaa   | 5460 |
|               | -            |              | JJ           |              | guaua        |      |

```
agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc 5520
tggtatcggt ctgcgattcc gactcgtcca acatcaatac aacctattaa tttcccctcg 5580
tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc cggtqaqaat 5640
ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca 5700
tcaaaatcac tcgcatcaac caaaccgtta ttcattcgtg attgcgcctg agcgagacga 5760
aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa ccggcgcagg 5820
aacactgcca gcgcatcaac aatattttca cctgaatcag gatattcttc taatacctgg 5880
aatgctgttt tcccggggat cgcagtggtg agtaaccatg catcatcagg agtacggata 5940
aaatgettga tggteggaag aggeataaat teegteagee agtttagtet gaecatetea 6000
tetgtaacat cattggcaac getacetttg ccatgtttca gaaacaacte tggcgcatcg 6060
ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc gcgagcccat 6120
ttatacccat ataaatcagc atccatgttg gaatttaatc gcggcctcga gcaagacgtt 6180
tcccgttgaa tatggctcat aacacccctt gtattactgt ttatgtaagc agacagtttt 6240
attgttcatg atgatatatt tttatcttgt gcaatgtaac atcagagatt ttgagacaca 6300
acgtggcttt ccccccccc ccattattga agcatttatc agggttattg tctcatgagc 6360
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 6420
cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat 6480
aggogtatea cgaggocott togto
```

```
<210> 16
<211> 6505
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 16
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tettatgeat getatactgt ttttggettg gggeetatae acceegett cettatgeta 1140
taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecce 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagaetgae aeggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctet tetecggtag eggeggaget tecacateeg ageeetggte ceatgeetee 1500
```

```
ageggeteat ggtegetegg eageteettg etectaacag tggaggecag acttaggeae 1560
 agcacaatge ccaccaccac cagtgtgeeg cacaaggeeg tggeggtagg gtatgtgtet 1620
 gaaaatgage gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcageg 1680
 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
 gtgaaggaga agtaccagca cetgtggcgc tggggctggc gctggggcac catgctgctg 1980
 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
 cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
 acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
 agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
 accagcatec geggeaaggt geagaaggag taegeettet tetacaaget ggacateate 2460
 cccatcgaca acgacaccac cagetacage etgaccaget gcaacacege tagcatcaac 2520
 tgcacccgcc ccaacaacaa cacccgcaag agcatccaca tcggccccgg ccgcgccttc 2580
 tacaccaccg gcgagatcat cggcgacatc cgccaggccc actgcaacct gagccgcgcc 2640
 aagtggaacg acaccctgaa caagatcgtg atcaagctgc gcgagcagtt cggcaacaag 2700
 accategtgt teaageacag cageggegge gacceegaga tegtgaceca cagetteaac 2760
 tgcggcggcg agttcttcta ctgcaacagc acccagctgt tcaacagcac ctggttcaac 2820
 agcacctgga gcaccgaggg cagcaacaac accgagggca gcgacaccat caccctgccc 2880
 tgccgcatca agcagatcat caacatgtgg cagaaggtgg gcaaggccat gtacgccccc 2940
 cccatcagcg gccagatccg ctgcagcagc aacatcaccg gcctgctgct gacccgcgac 3000
 ggcggcaaca gcaacaacga gagcgagate tteegteegg geggeggega catgegegae 3060
 aactggcgca gcgagctgta caagtacaag gtggtgaaga tcgagcccct gggcgtggcc 3120
 cccaccaagg ccaagettac cgtccaggcc cgccagctgc tgagcggcat cgtgcagcag 3180
 cagaacaacc tgctgcgcgc catcgaggcc cagcagcacc tgctgcagct gaccgtgtgg 3240
 ggcatcaagc agctgcaggc ccgcaccctg gccgtggagc gctacctgaa ggaccagcag 3300
 cagetegage agatetggaa ecacaceace tggatggagt gggacegega gateaacaac 3360
 tacaccagee tgatecacag cetgategag gagagecaga accageagga gaagaacgag 3420
caggagetge tggagetgga caagtgggee ageetgtgga actggttcaa catcaccaac 3480
tggctgtggt gaggatccag atctgctgtg ccttctagtt gccagccatc tgttgtttgc 3540
ccctccccg tgccttcctt gaccctggaa ggtgccactc ccactgtcct ttcctaataa 3600
gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcatgctgg ggatgcggtg 3720
ggctctatgg gtacccaggt gctgaagaat tgacccggtt cctcctgggc cagaaagaag 3780
cagging cacate controlling transaction to the carried transaction of the cagging that can be caused as the carried transaction of the carried transaction of
ccactcatag gacactcata gctcaggagg gctccgcctt caatcccacc cgctaaagta 3900
cttggagcgg tctctccctc cctcatcagc ccaccaaacc aaacctagcc tccaagagtg 3960
ggaagaaatt aaagcaagat aggctattaa gtgcagaggg agagaaaatg cctccaacat 4020
gtgaggaagt aatgagagaa atcatagaat tttaaggcca tcatggcctt aatcttccgc 4080
ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca 4140
ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 4200
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 4260
taggeteege eeceetgaeg ageateacaa aaategaege teaagteaga ggtggegaaa 4320
cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 4380
tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 4440
gettteteat ageteaeget gtaggtatet eagtteggtg taggtegtte geteeaaget 4500
gggctgtgtg cacgaaccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 4560
tettgagtec aacceggtaa gacacgaett ategecaetg geageageca etggtaacag 4620
gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta 4680
cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 4740
aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 4800
```

```
tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt 4860
ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag 4920
attatcaaaa aggatettea eetagateet tttaaattaa aaatgaagtt ttaaatcaat 4980
ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc 5040
tateteageg atetgtetat ttegtteate catagttgee tgaetegggg ggggggggeg 5100
ctgaggtctg cctcgtgaag aaggtgttgc tgactcatac caggcctgaa tcgccccatc 5160
atccagccag aaagtgaggg agccacggtt gatgagagct ttgttgtagg tggaccagtt 5220
ggtgattttg aacttttgct ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat 5280
etgateette aacteageaa aagttegatt tatteaacaa ageegeegte eegteaaqte 5340
agogtaatgo totgocagtg ttacaaccaa ttaaccaatt otgattagaa aaactcatog 5400
agcatcaaat gaaactgcaa tttattcata tcaggattat caataccata tttttgaaaa 5460
agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc 5520
tggtatcggt ctgcgattcc gactcgtcca acatcaatac aacctattaa tttcccctcg 5580
tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat 5640
ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca 5700
tcaaaatcac tcgcatcaac caaaccgtta ttcattcgtg attgcgcctg agcgagacga 5760
aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa ccggcgcagg 5820
aacactgcca gcgcatcaac aatattttca cctgaatcag gatattcttc taatacctgg 5880
aatgctgttt tcccggggat cgcagtggtg agtaaccatg catcatcagg agtacggata 5940
aaatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct gaccatctca 6000
tetgtaacat cattggcaac getacetttg ceatgtttca gaaacaacte tggcgcatcg 6060
ggetteccat acaategata gattgtegea cetgattgee egacattate gegageecat 6120
ttatacccat ataaatcagc atccatgttg gaatttaatc gcggcctcga gcaagacgtt 6180
tcccgttgaa tatggctcat aacacccctt gtattactgt ttatgtaagc agacagtttt 6240
attgttcatg atgatatatt tttatcttgt gcaatgtaac atcagagatt ttgagacaca 6300
acgtggcttt ccccccccc ccattattga agcatttatc agggttattg tctcatgagc 6360
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 6420
cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat 6480
aggogtatea egaggeeett tegte
                                                                  6505
```

```
<210> 17
<211> 6646
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 17
tegegegettt eggegatgae ggegaaaace teegacacat geageteeeg gagaeggtea 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
```

catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840

```
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080
 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140
 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200
 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
 tattggctat atgccaatac tetgteette agagaetgae aeggaetetg tatttttaea 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcetgtgg gaccagagee tgaageeetg cgtgaagetg 2280
accccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacage ctgaccaget gcaacaccag cgtgatcacc 2520
caggcotgec ccaaggtgag cttcgagccc atccccaacc actactgcgc ccccgccggc 2580
ttcgccatcc tgaagtgcaa ggacaagaag ttcaacggca agggcccctg caccaacgtg 2640
agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700
ggtagcctgg ccgaggagga ggtggtgatc cgcagcgcta acttcgccga caacgccaag 2760
gtgatcatcg tgcagctgaa cgagagcgtg gagatcaact gcacccgccc caacaacaac 2820
accegeaaga geatecaeat eggeeeegge egegeettet acaceaeegg egagateate 2880
ggcgacatec gecaggeeea etgeaacetg getagettet aetgeaacag cacceagetg 2940
ttcaacagca cctggttcaa cagcacctgg agcaccgagg gcagcaacaa caccgagggc 3000.
agegacacea teaccetgee etgeegeate aageagatea teaacatgtg geagaaggtg 3060
ggcaaggcca tgtacgcccc ccccatcagc ggccagatcc gctgcagcag caacatcacc 3120
ggcctgctgc tgacccgcga cggcggcaac agcaacaacg agagcgagat cttccgtccg 3180
ggcggcggcg acatgcgcga caactggcgc agcgagctgt acaagtacaa ggtggtgaag 3240
ategageeee tgggegtgge eeccaceaag gecaagetta eegteeagge eegeeagetg 3300
ctgagcggca tcgtgcagca gcagaacaac ctgctgcgcg ccatcgaggc ccagcagcac 3360
ctgctgcagc tgaccgtgtg gggcatcaag cagctgcagg cccgcaccct ggccgtggag 3420
tgggaccgcg agatcaacaa ctacaccagc ctgatccaca gcctgatcga ggagagccag 3540
aaccagcagg agaagaacga gcaggagctg ctggagctgg acaagtgggc cagcctgtgg 3600
aactggttca acatcaccaa ctggctgtgg tgaggatcca gatctgctgt gccttctagt 3660
tgccagccat ctgttgtttg cccctcccc gtgccttcct tgaccctgga aggtgccact 3720
cccactgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat 3780
tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc 3840
aggcatgctg gggatgcggt gggctctatg ggtacccagg tgctgaagaa ttgacccggt 3900
tectectggg ccagaaagaa gcaggcacat eccettetet gtgacacace etgtecacge 3960
ccctggttct tagttccagc cccactcata ggacactcat agctcaggag ggctccgcct 4020
tcaatcccac ccgctaaagt acttggagcg gtctctccct ccctcatcag cccaccaaac 4080
caaacctagc ctccaagagt gggaagaaat taaagcaaga taggctatta agtgcagagg 4140
```

```
gagagaaaat gcctccaaca tgtgaggaag taatgagaga aatcatagaa ttttaaggcc 4200
abcategeet taatetteeg etteeteget cactgacteg etgegetegg tegttegget 4260
geggegageg gtateagete aeteaaagge ggtaataegg ttateeacag aateagggga 4320
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaaggc 4380
cgcgttgctg gcgtttttcc ataggctccg ccccctgac gagcatcaca aaaatcgacg 4440
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 4500
aageteecte gtgegetete etgtteegae eetgeegett aeeggataee tgteegeett 4560
tetecetteg ggaagegtgg egetttetea tageteaege tgtaggtate teagtteggt 4620
qtaqqtcqtt cqctccaaqc tqqqctqtqt qcacqaaccc cccqttcaqc ccqaccqctq 4680
cyccttatec ggtaactate gtettgagte caaceggta agacacgaet tategecaet 4740
ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 4800
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta tctgcgctct 4860
getgaageea gttaeetteg gaaaaagagt tggtagetet tgateeggea aacaaaceae 4920
cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 4980
tcaagaagat cetttgatet titetaeggg gietgaeget eagtggaaeg aaaacteaeg 5040
ttaaqqqatt ttqqtcatqa qattatcaaa aaqqatcttc acctaqatcc ttttaaatta 5100
aaaatqaaqt tttaaatcaa tctaaaqtat atatgaqtaa acttggtctg acaqttacca 5160
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 5220
ctgactcggg ggggggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata 5280
ccaqqcctga atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5340
tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc 5400
gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca 5460
aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 5520
tetgattaga aaaaeteate gageateaaa tgaaaetgea atttatteat ateaggatta 5580
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 5640
thecatagga toggaaqate etggtategg tetgegatte egactegtee aacateaata 5700
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 5760
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 5820
ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 5880
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 5940
atcgaatgca accggcgcag gaacactgcc agcgcatcaa caatattttc acctgaatca 6000
ggatattett etaatacetg gaatgetgtt theeegggga tegeagtggt gagtaaceat 6060
gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6120
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6180
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6240
ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt ggaatttaat 6300
egeggeeteg ageaagaegt tteeegttga atatggetea taacacccet tgtattactg 6360
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6420
catcagagat titigagacac aacgtggctt tcccccccc cccattattg aagcatttat 6480
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6540
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 6600
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc
```

```
<210> 18
<211> 6670
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 18
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
```

| cagettgtet | gtaagcggat | gccgggagca  | gacaagcccg | tcagggcgcg  | tcagcgggtg | 120  |
|------------|------------|-------------|------------|-------------|------------|------|
| ccggcgggtg | teggggetgg | cttaactatg  | cggcatcaga | gcagattgta  | ctgagagtgc | 180  |
| accatatgcg | gtgtgaaata | ccgcacagat  | gcgtaaggag | aaaataccgc  | atcagattgg | 240  |
| Ctattggcca | ttgcatacgt | tgtatccata  | tcataatatg | tacatttata  | ttggctcatg | 300  |
| tccaacatta | ccgccatgtt | gacattgatt  | attgactagt | tattaatagt  | aatcaattac | 360  |
| ggggtcatta | gttcatagcc | catatatgga  | gttccgcgtt | acataactta  | cggtaaatgg | 420  |
| cccgcctggc | tgaccgccca | acgacccccg  | cccattgacg | tcaataatga  | cgtatgttcc | 480  |
| catagtaacg | ccaataggga | ctttccattg  | acgtcaatgg | gtggagtatt  | tacggtaaac | 540  |
| tgcccacttg | gcagtacatc | aagtgtatca  | tatgccaagt | acgcccccta  | ttgacgtcaa | 600  |
| tgacggtaaa | tggcccgcct | ggcattatgc  | ccagtacatg | accttatggg  | actttcctac | 660  |
| ttggcagtac | atctacgtat | tagtcatcgc  | tattaccatg | gtgatgcggt  | tttggcagta | 720  |
| catcaatggg | cgtggatagc | ggtttgactc  | acggggattt | ccaaqtctcc  | accccattga | 780  |
| cgtcaatggg | agtttgtttt | ggcaccaaaa  | tcaacgggac | tttccaaaat  | gtcgtaacaa | 840  |
| ctccgcccca | ttgacgcaaa | tgggcggtag  | gcgtgtacgg | toggaggtet  | atataagcag | 900  |
| agctcqttta | gtgaaccgtc | agategeetg  | gagacgccat | ccacactatt  | ttgacctcca | 960  |
| tagaagacac | cqqqaccqat | ccaqcctccq  | cggccgggaa | cogtocatto  | gaacgcggat | 1020 |
| tecceatace | aagagtgacg | taagtaccgc  | ctatagactc | tataggcaca  | ccctttaac  | 3080 |
| tettatecat | gctatactgt | ttttaactta  | gggcctatac | acccccactt  | ccttatecta | 1140 |
| taggtgatgg | tatagettag | cctataggtg  | tgggttattg | accattatto  | accacteee  | 1200 |
| tattootoac | gatactttcc | attactaatc  | cataacatgg | ctctttacca  | gaagtatata | 1260 |
| tattggctat | atoccaatac | tctatccttc  | agagactgac | acquacteta  | tatttttaa  | 1220 |
| ggatggggtc | ccatttatta | tttacaaatt  | cacatataca | acggactetg  | cacccccaca | 1320 |
| caceatttt  | attaaacata | acatagaata  | tccacgcgaa | tetegggtag  | atattagge  | 1380 |
| catagactet | tetecaataa | caacaaaact  | tccacatccg | acceteate   | gegeeegga  | 1440 |
| agcggctcat | aatcactcaa | cagetgaagee | ctcctaacag | tagaaaaaaa  | acttacee   | 1200 |
| agcacaatoc | ccaccaccac | cagteteeeg  | cacaaggccg | tagaagataa  | accuaggeae | 7200 |
| gaaaatgage | atagagatta | agetegeeg   | gctgacgcag | atgasaga    | grangegree | 1620 |
| acadaadaad | atacagacag | ctarattatt  | geegaegeag | acggaagact  | Laaggeageg | 1680 |
| attacaatac | tattaacaat | gaagagaat   | gtattctgat | aagagccaga  | ggtaactccc | 1740 |
| cacacaacaa | gecaacgge  | ggagggcagt  | gtagtctgag | tagtactegt  | rgcrgeegeg | 1800 |
| tacaatcaca | gacacaacag | catatastas  | aacagactgt | accetatocae | gggtettte  | 1860 |
| atasaaaa   | antaccacca | cetatagaaa  | gatatcgcgg | actaga      | caccatgege | 1920 |
| gegaaggaga | tastataaa  | cccgcggcgc  | tggggctggc | tesestate   | catgetgetg | 1980 |
| cccatataca | agazagaga  | cgccaccgag  | aagctgtggg | tgaccgtgta  | ctacggcgtg | 2040 |
| accasaataa | aggaggccac | caccaccetg  | ctctgcgcca | gcgacgccaa  | ggcctacgac | 2100 |
| accyaggige | tagtassat  | ggeeaeeeae  | gcctgcgtgc | ccaecgaece  | caacccccag | 2160 |
| gagguggugu | aggagaacga | gaccgagaac  | ttcgacatgt | ggaagaacga  | carggragag | 2220 |
| cagacgcacg | aggacaccac | cageetgtgg  | gaccagagcc | cgaagecetg  | cgtgaagctg | 2280 |
| acceceege  | gegtgageet | gaagtgeace  | gacctgaaga | acgacaccaa  | caccaacagc | 2340 |
| ageageggee | gcatgateat | ggagaagggc  | gagatcaaga | actgcagett  | caacatcagc | 2400 |
| accageacce | geggeaaggt | gcagaaggag  | tacgccttct | tctacaagct  | ggacatcatc | 2460 |
| cccaccgaca | acgacaccac | cagetaeage  | ctgaccagct | gcaacaccag  | cgtgatcacc | 2520 |
| thananhan  | ccaaggtgag | ettegageee  | atccccaacc | actactgcgc  | ccccgccggc | 2580 |
| ttegecatee | tgaagtgcaa | ggacaagaag  | ttcaacggca | agggcccctg  | caccaacgtg | 2640 |
| agcaccgtgc | agtgcaccca | cggcatccgc  | cccgtggtga | gcacccaget  | gctgctgaac | 2700 |
| ggtageetgg | ccgaggagga | ggtggtgatc  | cgcagcgcta | acttcgccga  | caacgccaag | 2760 |
| grgarcareg | tgcagctgaa | cgagagcgtg  | gagatcaact | gcacccgccc  | caacaacaac | 2820 |
| accegeaaga | gcatccacat | cggccccggc  | cgcgccttct | acaccaccgg  | cgagatcatc | 2880 |
| ggcgacatcc | gccaggccca | ctgcaacctg  | agccgcgcca | agtggaacga  | caccctgaac | 2940 |
| aagatcgtga | tcaagctgcg | cgagcagttc  | ggcaacaaga | ccatcgtgtt  | caagcacagc | 3000 |
| agcggcggcg | accccgagat | cgtgacccac  | agcttcaact | gcggcggcga  | gttcttctac | 3060 |
| tgcaacagca | cccagctgtt | caacagcacc  | tggttcaaca | gcacctggag  | caccgagggc | 3120 |
| agcaacaaca | ccgagggcag | cgacaccatc  | accctgccct | gccgcatcgc  | tagcagcaac | 3180 |
| aacgagagcg | agatetteeg | tccgggcggc  | ggcgacatgc | gcgacaactg  | gcgcagcgag | 3240 |
| ctgtacaagt | acaaggtggt | gaagatcgag  | cccctgggcg | tggcccccac  | caaggccaag | 3300 |
| cttaccgtcc | aggcccgcca | gctgctgagc  | ggcatcgtgc | agcagcagaa  | caacctgctg | 3360 |

|            |            |            |            |            |            | 2422 |
|------------|------------|------------|------------|------------|------------|------|
|            | aggcccagca |            |            |            |            |      |
|            | ccctggccgt |            |            |            |            |      |
| tggaaccaca | ccacctggat | ggagtgggac | cgcgagatca | acaactacac | cagcctgatc | 3540 |
| cacagcctga | tcgaggagag | ccagaaccag | caggagaaga | acgagcagga | gctgctggag | 3600 |
| ctggacaagt | gggccagcct | gtggaactgg | ttcaacatca | ccaactggct | gtggtgagga | 3660 |
|            | ctgtgccttc |            |            |            |            |      |
|            | tggaaggtgc |            |            |            |            |      |
|            | tgagtaggtg |            |            |            |            |      |
|            | gggaagacaa |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | agaattgacc |            |            |            |            |      |
|            | caccctgtcc |            |            |            |            |      |
|            | ggagggctcc |            |            |            |            |      |
|            | tcagcccacc |            |            |            |            |      |
| aagataggct | attaagtgca | gagggagaga | aaatgcctcc | aacatgtgag | gaagtaatga | 4200 |
| gagaaatcat | agaattttaa | ggccatcatg | gccttaatct | tccgcttcct | cgctcactga | 4260 |
| ctcgctgcgc | teggtegtte | ggctgcggcg | agcggtatca | gctcactcaa | aggcggtaat | 4320 |
|            | acagaatcag |            |            |            |            |      |
|            | aaccgtaaaa |            |            |            |            |      |
|            | cacaaaaatc |            |            |            |            |      |
|            | gcgtttcccc |            |            |            |            |      |
|            | tacctgtccg |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | tatctcagtt |            |            |            |            |      |
|            | cagcccgacc |            |            |            |            |      |
|            | gacttatcgc |            |            |            |            |      |
|            | ggtgctacag |            |            |            |            |      |
|            | ggtatctgcg |            |            |            |            |      |
|            | ggcaaacaaa |            |            |            |            |      |
|            | agaaaaaaag |            |            |            |            |      |
| cgctcagtgg | aacgaaaact | cacgttaagg | gattttggtc | atgagattat | caaaaaggat | 5100 |
| cttcacctag | atccttttaa | attaaaaatg | aagttttaaa | tcaatctaaa | gtatatatga | 5160 |
| gtaaacttgg | tctgacagtt | accaatgctt | aatcagtgag | gcacctatct | cagcgatctg | 5220 |
| tctatttcgt | tcatccatag | ttgcctgact | cggggggggg | gggcgctgag | gtctgcctcg | 5280 |
| tgaagaaggt | gttgctgact | cataccaggc | ctgaatcgcc | ccatcatcca | gccagaaagt | 5340 |
|            | cggttgatga |            |            |            |            |      |
|            | acggaacggt |            |            |            |            |      |
|            | cgatttattc |            |            |            |            |      |
|            | accaattaac |            |            |            |            |      |
|            | tcatatcagg |            |            |            |            |      |
|            | actcaccgag |            |            |            |            |      |
|            | gtccaacatc |            |            |            |            |      |
|            | aatcaccatg |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | agacttgttc |            |            |            |            |      |
|            | cgttattcat |            |            |            |            |      |
|            | aattacaaac |            |            |            |            |      |
|            | tttcacctga |            |            |            |            |      |
|            | tggtgagtaa |            |            |            |            |      |
|            | taaattccgt |            |            |            |            |      |
|            | ctttgccatg |            |            |            |            |      |
|            | tcgcacctga |            |            |            |            |      |
| tcagcatcca | tgttggaatt | taatcgcggc | ctcgagcaag | acgtttcccg | ttgaatatgg | 6360 |
| ctcataacac | cccttgtatt | actgtttatg | taagcagaca | gttttattgt | tcatgatgat | 6420 |
| atattttat  | cttgtgcaat | gtaacatcag | agattttgag | acacaacgtg | gctttccccc | 6480 |
|            | attgaagcat |            |            |            |            |      |
|            | aaaataaaca |            |            |            |            |      |
|            | aaaccattat |            |            |            |            |      |
|            |            |            |            |            |            |      |

ccctttcgtc 6670

<210> 19 <211> 6709 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 ageteqttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac egggacegat eeageeteeg eggeegggaa eggtgeattg gaacgeggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eeeetttgge 1080 tettatgeat getataetgt tittqqettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cetataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac totgtootto agagactgac acggactotg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccqt cccccgtgcc 1380 egeagttttt attaaacata gegtgggate tecaegegaa tetegggtae gtgtteegga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 aggggeteat ggtcqctcqq cageteettg etectaacag tggaggecag acttaggeac 1560 ageacaatge ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcaqaaqaag atgcaqqcaq ctgaqttgtt qtattctqat aaqaqtcaqa ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040 cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280 accecetgt gegtgageet gaagtgeace gacetgaaga acgacaceaa caccaacage 2340 agcageggee geatgateat ggagaaggge gagateaaga actgeagett caacateage 2400 accagcatoc goggcaaggt goagaaggag tacgcottot totacaagot ggacatcato 2460 cccategaca acgacaccae cagetacage etgaccaget geaacaccag egtgateace 2520

- 44 -

| caggettee  | ccaanntnan | cttccaccc  | atccccaacc | 2012010000  |             | 2500         |
|------------|------------|------------|------------|-------------|-------------|--------------|
| ttccccatcc | taaataaa   | ggiggggg   | ttassaaas  | accaccaca   | anagagaga   | 2500         |
| accaccatac | agtagagaga | ggacaagaag | ttcaacggca | agggeeeeg   | caccaacgtg  | 2640         |
| agtactgtgt | agegeaceea | eggeateege | cccgtggtga | geaeceaget  | gctgctgaac  | 2700         |
| ggcagcctgg | ccgaggagga | ggtggtgate | cgcagcgcta | acttcgccga  | caacgccaag  | 2760         |
| grgarcarcg | tgcagctgaa | cgagagcgtg | gagatcaact | gcacccgccc  | caacaacaac  | 2820         |
|            |            |            | cgcgccttct |             |             |              |
|            |            |            | agccgcgcca |             |             |              |
| aagatcgtga | tcaagctgcg | cgagcagttc | ggcaacaaga | ccatcgtgtt  | caagcacagc  | 3000         |
|            |            |            | agcttcaact |             |             |              |
| tgcaacagca | cccagctgtt | caacagcacc | tggttcaaca | gcacctggag  | caccgagggc  | 3120         |
| agcaacaaca | ccgagggcag | cgacaccatc | accctgccct | gccgcatcaa  | gcagatcatc  | 3180         |
| aacatgtggc | agaaggtggg | caaggccatg | tacgccccc  | ccatcagcgg  | ccagatccgc  | 3240         |
| tgcagcagca | acatcaccgg | cctgctgctg | acccgcgacg | gcggcaacag  | caacaacgag  | 3300         |
| agcgagatct | tccgtccggg | cggcggcgct | agcgccaagc | ttaccqtcca  | qqcccqccaq  | 3360         |
| ctgctgagcg | gcatcgtgca | gcagcagaac | aacctgctgc | gcgccatcga  | gacccagcag  | 3420         |
| cacctgctgc | agctgaccgt | gtggggcatc | aagcagctgc | aggcccgcac  | cctaaccata  | 3480         |
|            |            |            | gagcagatct |             |             |              |
| gagtgggacc | gcgagatcaa | caactacacc | agcctgatcc | acacctcat   | caecaggacg  | 3500         |
| cadaaccadc | addadaadaa | caaccacac  | ctgctggagc | tageetgae   | gaggagage   | 3660         |
| tagaactagt | tcaacatcac | caactaagta | tggtgaggat | casastataa  | ggccagcccg  | 3000         |
| agttaccage | cabatattat | ttacagetag | cggcgaggat | coagacetge  | rgrgeereer  | 3720         |
| agregeeage | tacttgccgc | atasastasa | cccgtgcctt | corregaceer | ggaaggtgee  | 3780         |
|            |            |            | gaaattgcat |             |             |              |
| cattetatte | r999999199 | ggrggggcag | gacagcaagg | gggaggattg  | ggaagacaat  | 3900         |
|            |            |            | atgggtaccc |             |             |              |
| ggttcctcct | gggccagaaa | gaagcaggca | catccccttc | tctgtgacac  | accctgtcca  | 4020         |
|            |            |            | ataggacact |             |             |              |
|            |            |            | geggtetete |             |             |              |
| aaccaaacct | agcctccaag | agtgggaaga | aattaaagca | agataggcta  | ttaagtgcag  | 4200         |
| agggagagaa | aatgcctcca | acatgtgagg | aagtaatgag | agaaatcata  | gaattttaag  | 4260         |
| gccatcatgg | ccttaatctt | ccgcttcctc | gctcactgac | tcgctgcgct  | cggtcgttcg  | 4320         |
| gctgcggcga | gcggtatcag | ctcactcaaa | ggcggtaata | cggttatcca  | cagaatcagg  | 4380         |
| ggataacgca | ggaaagaaca | tgtgagcaaa | aggccagcaa | aaggccagga  | accgtaaaaa  | 4440         |
| ggccgcgttg | ctggcgtttt | tccataggct | ccgccccct  | gacgagcatc  | acaaaaatcg  | 4500         |
| acgctcaagt | cagaggtggc | gaaacccgac | aggactataa | agataccagg  | cgtttcccc   | 4560         |
| tggaagctcc | ctcgtgcgct | ctcctgttcc | gaccctgccg | cttaccggat  | acctgtccgc  | 4620         |
| ctttctccct | tcgggaagcg | tggcgctttc | tcatagctca | cactatagat  | atctcagttc  | 4680         |
|            |            |            | tgtgcacgaa |             |             |              |
|            |            |            | gtccaacccg |             |             |              |
|            |            |            | cagagcgagg |             |             |              |
| gttcttgaag | taataaccta | actacoocta | cactagaaga | acagtatttg  | gtatctgcgc  | 4920         |
| tctgctgaag | ccagttacct | tcggaaaaaa | agttggtagc | tetteateea  | dcasacasac  | 4980         |
| caccactaat | agcggtggtt | ttttattta  | caagcagcag | attaccccc   | geadacaaac  | 5040         |
| atctcaagaa | gateetttea | tetttetae  | ggggtctgac | actacgegea  | aggaaaaaagg | 5040<br>E100 |
| acottaagaa | attttaatca | tagasttata | aaaaaggatc | ttanastasa  | tagtttt     | 2100         |
| ttaaaaatga | actitional | caatetaaae | tatatatgag | troacctaga  | ccccccaaa   | 2100         |
| ccaatactta | agecetaaae | caacccaaag | acacacyay  | caaacttggt  | ctgacagtta  | 5220         |
| tasatasata | accagigagg | cacctatete | agcgatctgt | ctatttegtt  | catecatage  | 5280         |
| change     | 999999999  | ggcgctgagg | tctgcctcgt | gaagaaggtg  | ttgctgactc  | 5340         |
| acaccaggee | tgaatcgccc | catcatccag | ccagaaagtg | agggagccac  | ggttgatgag  | 5400         |
| agettegttg | taggtggacc | agttggtgat | tttgaacttt | tgctttgcca  | cggaacggtc  | 5460         |
| egegetgteg | ggaagatgcg | tgatctgatc | cttcaactca | gcaaaagttc  | gatttattca  | 5520         |
| acaaagccgc | cgtcccgtca | agtcagcgta | atgctctgcc | agtgttacaa  | ccaattaacc  | 5580         |
|            |            |            | aaatgaaact |             |             |              |
| ttatcaatac | catatttttg | aaaaagccgt | ttctgtaatg | aaggagaaaa  | ctcaccgagg  | 5700         |
| cagttccata | ggatggcaag | atcctggtat | cggtctgcga | ttccgactcg  | tccaacatca  | 5760         |
| atacaaccta | ttaatttccc | ctcgtcaaaa | ataaggttat | caagtgagaa  | atcaccatga  | 5820         |
|            |            |            |            |             |             |              |

```
gtgacgactgaatccggtgagaatggcaaaagcttatgcatttctttccagacttgttca5880acaggccagccattacgctcgtcatcaaaatcactcgcatcaaccaaaccgttattcatt5940cgtgattgcgcctgagcgagacgaaatacgcgatcgctgttaaaaggacaattacaaaca6000ggaatcgaatgcaaccggcgcaggaacactgccagcgcatcaacaatattttcacctgaa6060tcaggatattcttctaatacctggaatgctgttttcccggggatcgcagtggtgagtaac6120catgcatcatcaggagtacggataaaatgcttgatggtcggaagaggcataaattccgtc6180agccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttgccatgt6240ttcagaaacaactctggcgcccattataaccagcatcatgttggaattt6300tgcccgacattatcgcgagcccatttataccagcatcatgttggaattt6360aatcgcggcctcgagcaagacgtttcccgttgaatatggctcataacaccccttgtatta6420ctgtttatgaagcagacagtttattgtcatgatgatatattttatcttggaattt6540tatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaa6600ataggggttccgcgcacatttccccgaaaagtgccacctgacttcgcccttcgcc6709
```

<210> 20 <211> 6736 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatage ggtttgacte acggggattt ccaagtetee accecattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagaetgae aeggaetetg tatttttaea 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

```
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accetgetet gegccagega egccaaggec tacgacaceg 2100
aggtgcacaa cgtgtgggcc acceacgcct gcgtgcccac cgaccccaac ccccaggagg 2160
tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280
ccctgtgcgt gagcctgaag tgcaccgacg ctagcaagaa ctgcagcttc aacatcagca 2340
ccagcatccg cggcaaggtg cagaaggagt acgccttctt ctacaagctg gacatcatcc 2400
ccatcgacaa cgacaccacc agctacagcc tgaccagctg caacaccagc gtgatcaccc 2460
aggeotycee caaggtgage ttegageeea teeccaacea ctactgegee eccqeeqqet 2520
tegecatect gaagtgcaag gacaagaagt teaacggcaa gggcccetgc accaacgtga 2580
gcaccgtgca gtgcacccac ggcatccgcc ccgtggtgag cacccagctg ctgctgaacg 2640
gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcgccgac aacgccaagg 2700
tgatcatcgt gcagctgaac gagagcgtgg agatcaactg cacccgcccc aacaacaaca 2760
cccgcaagag catccacatc ggccccggcc gcgccttcta caccaccggc gagatcatcg 2820
gcgacatecg ccaggcccac tgcaacetga gccgcgccaa gtggaacgac accctgaaca 2880
agategtgat caagetgege gageagtteg geaacaagae categtgtte aageacagea 2940
geggeggega cecegagate gtgacecaca getteaactg eggeggegag ttettetact 3000
gcaacagcac ccagctgttc aacagcacct ggttcaacag cacctggagc accgagggca 3060
acatgtggca gaaggtgggc aaggccatgt acgcccccc catcagcggc cagatccgct 3180
gcagcagcaa catcaccggc ctgctgctga cccgcgacgg cggcaacagc aacaacgaga 3240
gcgagatett ccgtccgggc ggcggcgaca tgcgcgacaa ctggcgcagc gagctgtaca 3300
agtacaaggt ggtgaagatc gagccctgg gcgtggcccc caccaaggcc aagcttaccg 3360
tecaggeecg ceagetgetg ageggeateg tgeageagea gaacaacetg etgegegeea 3420
tegaggecca geageacetg etgeagetga eegtgtgggg cateaageag etgeaggece 3480
gcaccctggc cgtggagcgc tacctgaagg accagcagca gctcgagcag atctggaacc 3540
acaccacctg gatggagtgg gaccgcgaga tcaacaacta caccagcctg atccacagcc 3600
tgatcgagga gagccagaac cagcaggaga agaacgagca ggagctgctg gagctggaca 3660
agtgggccag cctgtggaac tggttcaaca tcaccaactg gctgtggtga ggatccagat 3720
ctgctgtgcc ttctagttgc cagccatctg ttgtttgccc ctccccgtg ccttccttga 3780
ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt gcatcgcatt 3840
attgggaaga caatagcagg catgctgggg atgcggtggg ctctatgggt acccaggtgc 3960
tgaagaattg accoggitec teetgggeea gaaagaagea ggeacateee ettetetgtg 4020
acacaccetg tecaegeece tggttettag ttecageece acteatagga cacteatage 4080
traggagggc trogcottca atrocarceg ctaaagtact tggagcggtc totocotcoc 4140
teateagece accaaaceaa acetageete caagagtggg aagaaattaa ageaagatag 4200
gctattaagt gcagagggag agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat 4260
catagaattt taaggccatg atttaaggcc atcatggcct taatetteeg cttccteget 4320
cactgacteg etgegetegg tegttegget geggegageg gtateagete acteaaagge 4380
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 4440
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 4500
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 4560
actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac 4620
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 4680
tageteaege tgtaggtate teagtteggt gtaggtegtt egeteeaage tgggetgtgt 4740
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 4800
caacceggta agacacgact tategecact ggcagcagee actggtaaca ggattageag 4860
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 4920
tagaagaaca gtatttggta tetgegetet getgaageea gttacetteg gaaaaagagt 4980
```

```
tggtagetet tgateeggea aacaaaceae egetggtage ggtggttttt ttgtttgeaa 5040
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 5100
gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 5160
aaggatette acctagatee tittaaatta aaaatgaagt titaaateaa tetaaagtat 5220
atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 5280
gatctqtcta tttcgttcat ccataqttqc ctgactcqqq qqqqqqqqc gctgaggtct 5340
geotegtgaa gaaggtgttg etgacteata eeaggeetga ategeeceat catecageea 5400
gaaagtgagg gagccacggt tgatgagagc tttgttgtag gtggaccagt tggtgatttt 5460
gaacttttgc tttgccacgg aacggtctgc gttgtcggga agatgcgtga tctgatcctt 5520
caactcagca aaagttcgat ttattcaaca aagccgccgt cccgtcaagt cagcgtaatg 5580
etetgecagt gttacaacca attaaccaat tetgattaga aaaactcate gagcatcaaa 5640
tgaaactgca atttattcat atcaggatta tcaataccat atttttgaaa aagccgtttc 5700
tqtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg 5760
tctqcqattc cqactcqtcc aacatcaata caacctatta atttcccctc gtcaaaaata 5820
aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc 5880
ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca 5940
ctcgcatcaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg aaatacgcga 6000
togotqttaa aaggacaatt acaaacagga atogaatgca accggcgcag gaacactgcc 6060
agggcatcaa caatattttc acctgaatca ggatattctt ctaatacctg gaatgctgtt 6120
ttcccqqqqa tcqcaqtggt gagtaaccat gcatcatcag gagtacggat aaaatgcttg 6180
atgqtcqqaa qaqqcataaa ttccqtcaqc caqtttaqtc tqaccatctc atctqtaaca 6240
tcattqqcaa cqctaccttt gccatgtttc agaaacaact ctgqcgcatc qggcttccca 6300
tacaategat agattgtege acetgattge eegacattat egegageeca tttataccca 6360
tataaatcag catccatgtt ggaatttaat cgcggcctcg agcaagacgt ttcccgttga 6420
atatggctca taacacccct tgtattactg tttatgtaag cagacagttt tattgttcat 6480
qatqatatat ttttatcttg tqcaatgtaa catcagagat tttgagacac aacgtggctt 6540
teceeece eccattattg aagcatttat cagggttatt gteteatgag eggatacata 6600
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 6660
ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc 6720
acgaggeeet ttegte
```

```
<210> 21
<211> 6697
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegttt eggtgatgae ggtgaaaace tetgaeacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
qqqqtcatta gttcataqcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccqcctqgc tgaccqccca acqaccccq cccattgacg tcaataatqa cqtatqttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
```

|            |            | 4          |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | agtttgttt  |            |            |            |            |      |
|            | ttgacgcaaa |            |            |            |            |      |
|            | gtgaaccgtc |            |            |            |            |      |
|            | cgggaccgat |            |            |            |            |      |
|            | aagagtgacg |            |            |            |            |      |
|            | gctatactgt |            |            |            |            |      |
|            | tatagcttag |            |            |            |            |      |
|            | gatactttcc |            |            |            |            |      |
|            | atgccaatac |            |            |            |            |      |
|            | ccatttatta |            |            |            |            |      |
|            | attaaacata |            |            |            |            |      |
|            | tctccggtag |            |            |            |            |      |
|            | ggtcgctcgg |            |            |            |            |      |
|            | ccaccaccac |            |            |            |            |      |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact | taaggcagcg | 1680 |
|            | atgcaggcag |            |            |            |            |      |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg | 1800 |
|            | gacataatag |            |            |            |            |      |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | atgcgcgtga | 1920 |
| aggagaagta | ccagcacctg | tggcgctggg | gctggcgctg | gggcaccatg | ctgctgggca | 1980 |
|            | ctgcagcgcc |            |            |            |            |      |
| tgtggaagga | ggccaccacc | accctgctct | gcgccagcga | cgccaaggcc | tacgacaccg | 2100 |
| aggtgcacaa | cgtgtgggcc | acccacgcct | gcgtgcccac | cgaccccaac | ccccaggagg | 2160 |
| tggtgctggt | gaacgtgacc | gagaacttcg | acatgtggaa | gaacgacatg | gtggagcaga | 2220 |
| tgcacgagga | catcatcagc | ctgtgggacc | agagcctgaa | gccctgcgtg | aagctgaccc | 2280 |
| ccctgtgcgt | gagcctgaag | tgcaccgacc | tgaagaacga | caccaacacc | aacagcagca | 2340 |
| gcggccgcat | gatcatggag | aagggcgaga | tcaagaactg | cagcttcgct | agcaccagct | 2400 |
| gcaacaccag | cgtgatcacc | caggcctgcc | ccaaggtgag | cttcgagccc | atccccaacc | 2460 |
|            | ccccgccggc |            |            |            |            |      |
|            | caccaacgtg |            |            |            |            |      |
| gcacccagct | gctgctgaac | ggtagcctgg | ccgaggagga | ggtggtgatc | cgcagcgcta | 2640 |
| acttcgccga | caacgccaag | gtgatcatcg | tgcagctgaa | cgagagcgtg | gagatcaact | 2700 |
| gcacccgccc | caacaacaac | acccgcaaga | gcatccacat | cggccccggc | cgcgccttct | 2760 |
| acaccaccgg | cgagatcatc | ggcgacatcc | gccaggccca | ctgcaacctg | agccgcgcca | 2820 |
|            | caccctgaac |            |            |            |            |      |
| ccatcgtgtt | caagcacagc | agcggcggcg | accccgagat | cgtgacccac | agcttcaact | 2940 |
| gcggcggcga | gttcttctac | tgcaacagca | cccagctgtt | caacagcacc | tggttcaaca | 3000 |
|            | caccgagggc |            |            |            |            |      |
| gccgcatcaa | gcagatcatc | aacatgtggc | agaaggtggg | caaggccatg | tacgcccccc | 3120 |
| ccatcagcgg | ccagatccgc | tgcagcagca | acatcaccgg | cctgctgctg | acccgcgacg | 3180 |
| gcggcaacag | caacaacgag | agcgagatct | tccgtccggg | cggcggcgac | atgcgcgaca | 3240 |
|            | cgagctgtac |            |            |            |            |      |
| ccaccaaggc | caagcttacc | gtccaggccc | gccagctgct | gagcggcatc | gtgcagcagc | 3360 |
| agaacaacct | gctgcgcgcc | atcgaggccc | agcagcacct | gctgcagctg | accgtgtggg | 3420 |
|            | gctgcaggcc |            |            |            |            |      |
|            | gatctggaac |            |            |            |            |      |
|            | gatccacagc |            |            |            |            |      |
| aggagctgct | ggagctggac | aagtgggcca | gcctgtggaa | ctggttcaac | atcaccaact | 3660 |
|            | aggatccaga |            |            |            |            |      |
|            | gccttccttg |            |            |            |            |      |
|            | tgcatcgcat |            |            |            |            |      |
|            | caagggggag |            |            |            |            |      |
|            | tacccaggtg |            |            |            |            |      |
|            | ccttctctgt |            |            |            |            |      |
|            | acactcatag |            |            |            |            |      |
|            |            |            |            |            | •          |      |

```
ttggageggt ctctccctcc ctcatcagec caccaaacca aacctagect ccaagagtgg 4140
gaagaaatta aagcaagata ggctattaag tgcagaggga gagaaaatgc ctccaacatg 4200
tgaggaagta atgagagaaa tcatagaatt ttaaggccat gatttaaggc catcatggcc 4260
ttaatcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc 4320
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4380
aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 4440
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4500
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4560
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 4620
gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcaqttcqq tqtagqtcqt 4680
tegetecaag etgggetgtg tgeacgaace eccegtteag eccgaceget gegeettate 4740
cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc 4800
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 4860
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 4920
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4980
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 5040
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 5100
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5160
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 5220
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5280
999999999 cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat accaggcctg 5340
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttgttgta 5400
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcqqq 5460
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5520
tecegteaag teagegtaat getetgeeag tgttacaace aattaaceaa ttetgattag 5580
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5640
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5700
atggcaagat cotggtatcg gtotgcgatt cogactogto caacatcaat acaacctatt 5760
aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5820
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5880
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 5940
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 6000
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 6060
tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tgcatcatca 6120
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6180
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6240
tetggegeat egggetteec atacaatega tagattgteg cacetgattg ecegacatta 6300
tegegagece atttatacce atataaatea geatecatgt tggaatttaa tegeggeete 6360
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6420
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6480
ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta tcagggttat 6540
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6600
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6660
acctataaaa ataggcgtat cacgaggccc tttcgtc
                                                                  6697
```

```
<210> 22
<211> 6712
<212> DNA
<213> Artificial Sequence
<220>
```

<223> plasmid pVR1012x/s containing HIV genes

<400> 22

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaaccgtc agategcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccathtatta thtacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etcetaacag tggaggecag aettaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcq 1800 cqcgccacca gacataatag ctgacagact aacagactgt tcctttccat qqqtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920 aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040 tgtggaagga ggccaccacc accetgctct gegccagega egccaaggce tacgacaceg 2100 aggtgcacaa cgtgtgggcc acccacgcet gcgtgcccac cgaccccaac ccccaggagg 2160 tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220 tgcacgagga catcatcagc ctgtgggacc agagcetgaa gccetgcgtg aagetgaccc 2280 ccctgtgcgt gagcctgaag tgcaccgacc tgaagaacga caccaacacc aacagcagca 2340 gcggccgcat gatcatggag aagggcgaga tcaagaactg cagcttcaac atcagcacca 2400 geateegegg caaggtgeag aaggagtaeg cettetteta caagetggae ateateecea 2460 tegacaacga caccaceage tacageetga ecagetgeaa caccagegtg atcacceagg 2520 cetgeeccaa ggtgagette gageceatec ceaaceacta etgegeecce geeggetteg 2580 ccatcctgaa gtgcaaggac aagaagttca acggcaaggg cccttgcacc aacgtgagca 2640 ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg ctgaacggta 2700 gcctggccga ggaggaggtg gtgatccgca gcgctaactt cgccgacaac gccaaggtga 2760 tcatcgtgca gctgaacgag agcgtggaga tcaactgcac ccgcgctagc gcccactgca 2820 acctgagccg cgccaagtgg aacgacaccc tgaacaagat cgtgatcaag ctgcgcgagc 2880 agtteggeaa caagaceate gtgtteaage acageagegg eggegaeeee gagategtga 2940 eccacagett caactgegge ggegagttet tetactgeaa cageacceag etqtteaaca 3000 gcacctggtt caacagcacc tggagcaccg agggcagcaa caacaccgag qqcaqcgaca 3060 ccatcaccct gccctgccgc atcaagcaga tcatcaacat gtggcagaag gtgggcaagg 3120 ccatgtacgc ccccccatc ageggccaga tccgctgcag cagcaacatc accggcctgc 3180 tgctgacccg cgacggcggc aacagcaaca acgagagcga gatcttccgt ccgggcggcg 3240

| WO 02/3294. | 5          |            |            |             | PCT/US     | 301/2572 |
|-------------|------------|------------|------------|-------------|------------|----------|
| gcgacatgcg  | cgacaactgg | cgcagcgagc | tgtacaagta | caaggtggtg  | aagatcgagc | 3300     |
| ccctgggcgt  | ggcccccacc | aaggccaagc | ttaccgtcca | ggcccgccag  | ctgctgagcg | 3360     |
| gcatcgtgca  | gcagcagaac | aacctgctgc | gcgccatcga | ggcccagcag  | cacctgctgc | 3420     |
| agctgaccgt  | gtggggcatc | aagcagctgc | aggcccgcac | cctggccgtg  | gagcgctacc | 3480     |
| tgaaggacca  | gcagcagctc | gagcagatct | ggaaccacac | cacctggatg  | gagtgggacc | 3540     |
| gcgagatcaa  | caactacacc | agcctgatcc | acagectgat | cgaggagagc  | cagaaccagc | 3600     |
| aggagaagaa  | cgagcaggag | ctgctggagc | tggacaagtg | ggccagcctg  | tggaactggt | 3660     |
| tcaacatcac  | caactggctg | tggtgaggat | ccagatctgc | tgtgccttct  | agttgccagc | 3720     |
| catctgttgt  | ttgcccctcc | cccgtgcctt | ccttgaccct | qqaaqqtqcc  | actcccactg | 3780     |
| tcctttccta  | ataaaatgag | gaaattgcat | cgcattgtct | gagtaggtgt  | cattctattc | 3840     |
| tggggggtgg  | ggtggggcag | gacagcaagg | gggaggattg | qqaaqacaat  | agcaggcatg | 3900     |
| ctggggatgc  | ggtgggctct | atgggtaccc | aggtgctgaa | gaattgaccc  | ggttcctcct | 3960     |
| gggccagaaa  | gaagcaggca | catccccttc | tetatgacac | accetateca  | cgcccctggt | 4020     |
| tcttagttcc  | agccccactc | ataggacact | catageteag | gagggctccg  | ccttcaatcc | 4080     |
| cacccgctaa  | agtacttgga | gcggtctctc | cctccctcat | cagcccacca  | aaccaaacct | 4140     |
| agcctccaag  | agtgggaaga | aattaaagca | agataggeta | ttaaqtqcaq  | agggagagaa | 4200     |
| aatgcctcca  | acatgtgagg | aagtaatgag | agaaatcata | gaattttaag  | gccatgattt | 4260     |
| aaggccatca  | tggccttaat | cttccqcttc | ctcactcact | gactcactac  | gctcggtcgt | 4320     |
| teggetgegg  | cgagcggtat | cageteacte | aaaggcggta | atacogttat  | ccacagaatc | 4380     |
| aggggataac  | qcaqqaaaqa | acatotoacc | aaaaggccag | caaaaaggcca | ggaaccgtaa | 4440     |
| aaaggccgcg  | ttactaacat | ttttccatag | actccaccc  | cctgacgage  | atcacaaaaa | 4500     |
| togacgotoa  | agtcagaggt | qqcqaaaccc | gacaggacta | taaagatacc  | aggcgtttcc | 4560     |
| ccctggaage  | teceteatae | actetectat | tccgaccctg | ccacttacca  | gatacctgtc | 4620     |
| cgcctttctc  | ccttcqqqaa | gcataacact | ttctcatage | tcacactata  | ggtatctcag | 4680     |
| ttcqqtqtaq  | qtcqttcqct | ccaagctggg | ctatatacac | dagcccccc   | ttcagcccga | 4740     |
| ccgctgcgcc  | ttatccqqta | actatogtot | tgagtccaac | ccaataaaac  | acgacttatc | 4800     |
| gccactggca  | gcagccactq | gtaacaggat | tagcagagcg | aggtatgtag  | gcggtgctac | 4860     |
| agagttcttg  | aagtggtggc | ctaactacqq | ctacactaga | agaacagtat  | ttggtatctg | 4920     |
| cgctctgctg  | aagccagtta | ccttcggaaa | aagagttagt | agctcttgat  | ccggcaaaca | 4980     |
| aaccaccgct  | ggtagcggtg | gttttttat  | ttqcaaqcaq | cagattacgc  | gcagaaaaaa | 5040     |
| aggatctcaa  | gaagatcctt | tgatcttttc | tacqqqqtct | gacgctcagt  | ggaacgaaaa | 5100     |
| ctcacgttaa  | gggattttgg | tcatgagatt | atcaaaaagg | atcttcacct  | agateetttt | 5160     |
| aaattaaaaa  | tgaagtttta | aatcaatcta | aagtatatat | gagtaaactt  | ggtctgacag | 5220     |
| ttaccaatgc  | ttaatcagtg | aggcacctat | ctcagcgatc | totctatttc  | gttcatccat | 5280     |
| agttgcctga  | ctcggggggg | gggggcgctg | aggtctgcct | cqtqaaqaaq  | gtgttgctga | 5340     |
| ctcataccag  | gcctgaatcg | ccccatcatc | cagccagaaa | qtqaqqaqc   | cacggttgat | 5400     |
| gagagetttg  | ttgtaggtgg | accagttggt | gattttgaac | ttttqctttq  | ccacggaacg | 5460     |
| gtctgcgttg  | tcgggaagat | gcgtgatctg | atccttcaac | tcagcaaaag  | ttcgatttat | 5520     |
| tcaacaaagc  | cgccgtcccg | tcaagtcagc | gtaatgetet | gccagtgtta  | caaccaatta | 5580     |
| accaattctg  | attagaaaaa | ctcatcgagc | atcaaatgaa | actqcaattt  | attcatatca | 5640     |
| ggattatcaa  | taccatattt | ttgaaaaagc | cgtttctgta | atgaaggaga  | aaactcaccg | 5700     |
| aggcagttcc  | ataggatggc | aagatcctgg | tateggtetg | cgattccgac  | tcgtccaaca | 5760     |
| tcaatacaac  | ctattaattt | cccctcgtca | aaaataaggt | tatcaagtga  | gaaatcacca | 5820     |
| tgagtgacga  | ctgaatccgg | tgagaatggc | aaaagcttat | gcatttcttt  | ccagacttgt | 5880     |
| tcaacaggcc  | agccattacg | ctcgtcatca | aaatcactcq | catcaaccaa  | accgttattc | 5940     |
| attcgtgatt  | gcgcctgagc | gagacgaaat | acgcgatcgc | tgttaaaagg  | acaattacaa | 6000     |
| acaggaatcg  | aatgcaaccg | gcgcaggaac | actgccagcg | catcaacaat  | attttcacct | 6060     |
| gaatcaggat  | attcttctaa | tacctggaat | gctgttttcc | cggggatcgc  | agtggtgagt | 6120     |
| aaccatgcat  | catcaggagt | acggataaaa | tgcttgatgg | tcqqaaqaqq  | cataaattcc | 6180     |
| gtcagccagt  | ttagtctgac | catctcatct | gtaacatcat | tggcaacgct  | acctttgcca | 6240     |
| tgtttcagaa  | acaactctgg | cgcatcgggc | ttcccataca | atcgatagat  | tgtcgcacct | 6300     |
| gattgcccga  | cattatcgcg | agcccattta | tacccatata | aatcagcatc  | catgttggaa | 6360     |
| tttaatcgcg  | gcctcgagca | agacgtttcc | cgttgaatat | qqctcataac  | accccttgta | 6420     |
| ttactgttta  | tgtaagcaga | cagttttatt | gttcatgatg | atatatttt   | atcttgtgca | 6480     |
| atgtaacatc  | agagatttta | agacacaacq | tggctttccc | CCCCCCCCC   | ttattgaagc | 6540     |
|             |            |            |            |             |            |          |

```
atttatcagg gttattgtct catgagegga tacatatttg aatgtattta gaaaaataaa 6600 caaatagggg tteegegeac attteeeega aaagtgeeac etgaegteta agaaaceatt 6660 attateatga cattaaceta taaaaatagg egtateaega ggeeettteg te 6712
```

```
<210> 23
<211> 6715
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tettatgeat getatactgt ttttggettg gggeetatac acceeegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgac aeggactetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg cageteettg etectaacag tggaggeeag acttaggeac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tqctqccqcq 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accetgetet gegccagega egccaaggec tacqacaccq 2100
aggtgcacaa cgtgtgggcc acceacgcct gcgtgcccac cgaccccaac ccccaggagg 2160
tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcage etgtgggace agageetgaa geeetgegtg aagetgacee 2280
ccctgtgcgt gagcctgaag tgcaccgacc tgaagaacga caccaacacc aacaqcaqca 2340
geggeegeat gateatggag aagggegaga teaagaaetg eagetteaae ateageaeca 2400
```

|            | 4          |            |            | _          |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | caaggtgcag |            |            |            |            |      |
|            | caccaccagc |            |            |            |            |      |
|            | ggtgagcttc |            |            |            |            |      |
|            | gtgcaaggac |            |            |            |            |      |
|            | cacccacggc |            |            |            |            |      |
|            | ggaggaggtg |            |            |            |            |      |
|            | gctgaacgag |            |            |            |            |      |
|            | ccacatcggc |            |            |            |            |      |
|            | ggcccactgc |            |            |            |            |      |
|            | gctgcgcgag |            |            |            |            |      |
|            | cgagatcgtg |            |            |            |            |      |
|            | cctgccctgc |            |            |            |            |      |
|            | egececece  |            |            |            |            |      |
|            | ccgcgacggc |            |            |            |            |      |
|            | gcgcgacaac |            |            |            |            |      |
|            | cataacccc  |            |            |            |            |      |
|            | gcagcagcag |            |            |            |            |      |
|            | cgtgtggggc |            |            |            |            |      |
|            | ccagcagcag |            |            |            |            |      |
|            | caacaactac |            |            |            |            |      |
|            | gaacgagcag |            |            |            |            |      |
|            | caccaactgg |            |            |            |            |      |
|            | tgtttgcccc |            |            |            |            |      |
|            | ctaataaaat |            |            |            |            |      |
|            | tggggtgggg |            |            |            |            |      |
|            | tgcggtgggc |            |            |            |            |      |
|            | aaagaagcag |            |            |            |            |      |
|            | tccagcccca |            |            |            |            |      |
|            | taaagtactt |            |            |            |            |      |
|            | aagagtggga |            |            |            |            |      |
|            | ccaacatgtg |            |            |            |            |      |
|            | tcatggcctt |            |            |            |            |      |
|            | cggcgagcgg |            |            |            |            |      |
|            | aacgcaggaa |            |            |            |            |      |
|            | gcgttgctgg |            |            |            |            |      |
|            | tcaagtcaga |            |            |            |            |      |
|            | agctccctcg |            |            |            |            |      |
|            | ctcccttcgg |            |            |            |            |      |
|            | taggtcgttc |            |            |            |            |      |
|            | gccttatccg |            |            |            |            |      |
|            | gcagcagcca |            |            |            |            |      |
|            | ttgaagtggt |            |            |            |            |      |
|            | ctgaagccag |            |            |            |            |      |
|            | gctggtagcg |            |            |            |            |      |
|            | caagaagatc |            |            |            |            |      |
|            | taagggattt |            |            |            |            |      |
|            | aaatgaagtt |            |            |            |            |      |
|            | tgcttaatca |            |            |            |            |      |
|            | tgactcgggg |            |            |            |            |      |
|            | caggcctgaa |            |            |            |            |      |
|            | ttgttgtagg |            |            |            |            |      |
|            | ttgtcgggaa |            |            |            |            |      |
|            | agccgccgtc |            |            |            |            |      |
|            | ctgattagaa |            |            |            |            |      |
| tcaggattat | caataccata | tttttgaaaa | agccgtttct | gtaatgaagg | agaaaactca | 5700 |

```
ccgaggcagt tccataggat ggcaagatcc tggtatcggt ctgcgattcc gactcgtcca 5760
acatcaatac aacctattaa tttcccctcg tcaaaaataa ggttatcaag tgagaaatca 5820
ccatgagtga cgactgaatc cggtgagaat ggcaaaaqct tatqcatttc tttccaqact 5880
tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac caaaccgtta 5940
ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa aggacaatta 6000
caaacaggaa tcgaatgcaa ccggcgcagg aacactgcca gcgcatcaac aatattttca 6060
cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat cgcagtggtg 6120
agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag aggcataaat 6180
teegteagee agtttagtet gaccatetea tetgtaacat cattggcaac getacetttq 6240
ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata gattgtcgca 6300
cetgattgcc egacattate gegageeeat ttataceeat ataaateage atceatgttg 6360
gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat aacaccctt 6420
gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt tttatcttgt 6480
gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc ccattattga 6540
agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 6600
aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 6660
attattatca tgacattaac ctataaaaat aggcgtatca cgaqqcctt tcqtc
<210> 24
<211> 6619
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 24
tegegegttt eggtgatgae ggtgaaaace tetgaeacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tecaacatta cegecatgit gacatigati attgactagi tattaatagi aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgaccccg cccattgacg tcaataatga cqtatqttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac totgtcottc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccqtqcc 1380
```

cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560

| agcaca        | atac   | ccaccaccac | cagtgtgccg   | cacaaggccg              | tggcggtagg   | gtatgtgtct   | 1620 |
|---------------|--------|------------|--------------|-------------------------|--------------|--------------|------|
| gaaaat        | caaac  | gtggagattg | ggctcgcacg   | gctgacgcag              | atggaagact   | taaggcagcg   | 1680 |
| gcagaa        | agaag  | atgcaggcag | ctgagttgtt   | gtattctgat              | aagagtcaga   | ggtaactccc   | 1740 |
| attac         | ratac  | tattaacagt | agagggcagt   | gtagtctgag              | cagtactcgt   | tgctgccgcg   | 1800 |
| Secari        | racca  | gacataatag | ctgacagact   | aacagactgt              | tectttccat   | gggtcttttc   | 1860 |
| tacaal        | rcacc  | greateasea | catataatca   | gatatcgcgg              | ccqctctaga   | atgcgcgtga   | 1920 |
| agrag.        | adta   | ccaccaccta | taacactaaa   | actagcacta              | gggcaccatg   | ctgctgggca   | 1980 |
| tacta         | ataat  | ctagaacaca | accaagaagc   | tatagataac              | cqtqtactac   | ggcgtgcccg   | 2040 |
| tataa         | 20940  | ccaccacc   | accetactet   | acaccaacaa              | caccaaaacc   | tacgacaccg   | 2100 |
| cgcgg         | 2222   | catatacacc | acceacacet   | acatacceae              | cgaccccaac   | ccccaggagg   | 2160 |
| agguge        | cacaa  | assatasco  | gagaacttcg   | acatgtggaa              | gaacgacatg   | gtggagcaga   | 2220 |
| tggtg         | cegge  | gaacgtgacc | ctataggacc   | agaggggaa               | accetacata   | aagctgaccc   | 2280 |
| tgcac         | gagga  | gaggatgaag | tacaccaaca   | charcaccad              | ctgcaacacc   | agcgtgatca   | 2340 |
| eeetg         |        | gageeegaag | agettegacg   | ccatccccaa              | ccactactac   | gccccgccg    | 2400 |
| cccag         | geerg  | ccccaaggcg | ageceegage   | agttcaacgg              | caacaacage   | tgcaccaacg   | 2460 |
| gette         | gecat  | acastagage | aaggacaaga   | acceptant               | dadcacccad   | ctactactaa   | 2520 |
| rgagc.        | accgt  | geagtgeace | cacggcatce   | tecacage                | taacttcocc   | ctgctgctga   | 2580 |
| acggt         | agcct  | ggccgaggag | gaggtggtga   | tecgeagege              | chacecegee   | gacaacgcca   | 2640 |
| aggtg         | atcat  | cgtgcagetg | aacgagagcg   | regagaccaa              | atacacacac   | cccaacaaca   | 2700 |
| acacc         | cgcaa  | gagcatccac | ateggeeeg    | geegegeett              | ccacaccacc   | ggcgagatca   | 2760 |
| tcggc         | gacat  | ccgccaggcc | cactgcaacc   | tgageegege              | caagtggaac   | gacaccctga   | 2020 |
| acaag         | atcgt  | gatcaagctg | cgcgagcagt   | teggeaacaa              | gaccatcgtg   | ttcaagcaca   | 2020 |
| gcagc         | ggcgg  | cgaccccgag | atcgtgaccc   | acagetteaa              | etgeggegge   | gagttcttct   | 2000 |
| actgc         | aacag  | cacccagctg | ttcaacagca   | cctggttcaa              | cagcacctgg   | agcaccgagg   | 2000 |
| gcagc         | aacaa  | caccgagggc | agcgacacca   | tcaccctgcc              | ctgccgcatc   | aagcagatca   | 3000 |
| tcaac         | atgtg  | gcagaaggtg | ggcaaggcca   | tgtacgcccc              | ccccatcagc   | ggccagatcc   | 3000 |
| gctgc         | agcag  | caacatcacc | ggcctgctgc   | tgacccgcga              | cggcggcaac   | agcaacaacg   | 3120 |
| agago         | gagat  | cttccgtccg | ggcggcggcg   | acatgcgcga              | caactggcgc   | agcgagctgt   | 3180 |
| acaag         | tacaa  | ggtggtgaag | atcgagcccc   | tgggcgtggc              | ccccaccaag   | gccaagctta   | 3240 |
| ccgtc         | caggc  | ccgccagctg | ctgagcggca   | tcgtgcagca              | gcagaacaac   | ctgctgcgcg   | 3300 |
| ccatc         | gaggc  | ccagcagcac | ctgctgcagc   | tgaccgtgtg              | gggcatcaag   | cagctgcagg   | 3360 |
| cccgc         | accct  | ggccgtggag | cgctacctga   | aggaccagca              | gcagctcgag   | cagatetgga   | 3420 |
| accac         | accac  | ctggatggag | tgggaccgcg   | agatcaacaa              | ctacaccagc   | ctgatccaca   | 3480 |
| gcctg         | gatoga | ggagagccag | aaccagcagg   | agaagaacga              | gcaggagctg   | ctggagctgg   | 3540 |
| acaag         | tgggc  | cagcctgtgg | aactggttca   | acatcaccaa              | ctggctgtgg   | tgaggatcca   | 3600 |
| gatct         | gctgt  | gccttctagt | tgccagccat   | ctgttgtttg              | cccctcccc    | gtgccttcct   | 3660 |
| tgaco         | ctqqa  | aggtgccact | cccactgtcc   | tttcctaata              | aaatgaggaa   | attgcatcgc   | 3720 |
| attgt         | ctgag  | taggtgtcat | tctattctgg   | ggggtggggt              | ggggcaggac   | agcaaggggg   | 3780 |
| aggat         | tggga  | agacaatagc | aggcatgctg   | gggatgcggt              | gggctctatg   | ggtacccagg   | 3840 |
| tacto         | gaagaa | ttgacccggt | tectectggg   | ccagaaagaa              | gcaggcacat   | ccccttctct   | 3900 |
| ataad         | cacacc | ctgtccacgo | ccctggttct   | tagttccagc              | cccactcata   | ggacactcat   | 3960 |
| agete         | aggag  | ggctccgcct | tcaatcccac   | ccgctaaagt              | acttggagcg   | gteteteet    | 4020 |
| ccctc         | catcag | cccaccaaac | caaacctage   | ctccaagagt              | gggaagaaat   | taaagcaaga   | 4080 |
| tagge         | ctatta | agtgcagagg | gagagaaaat   | gcctccaaca              | tgtgaggaag   | taatgagaga   | 4140 |
| aatca         | ataqaa | ttttaaggco | atgatttaag   | gccatcatgg              | ccttaatctt   | ccgcttcctc   | 4200 |
| actca         | actgac | tegetgeget | cggtcgttcg   | gctgcggcga              | gcggtatcag   | ctcactcaaa   | 4260 |
| aacaa         | ataata | cggttatcca | cagaatcagg   | <sub>I</sub> ggataacgca | . ggaaagaaca | tgtgagcaaa   | 4320 |
| agge          | caqcaa | aaggccagga | accgtaaaaa   | ggccgcgttg              | ctggcgtttt   | tccataggct   | 4380 |
| ccac          | cccct  | gacqaqcatc | : acaaaaatcg | , acgctcaagt            | . cagaggtggc | : gaaacccgac | 4440 |
| agga          | ctataa | agataccago | cgtttccccc   | : tggaagctcc            | : ctcgtgcgct | ctcctgttcc   | 4500 |
| gacco         | ctacco | cttaccqqat | acctatccac   | : ctttctccct            | tegggaageg   | tggcgctttc   | 4560 |
| tcata         | agetea | cqctataaat | atctcagttc   | ggtgtaggtc              | gttcgctcca   | agctgggctg   | 4620 |
| tata          | cacdaa | cccccatte  | agecegaces   | ctgcgcctta              | tccggtaact   | atcgtcttga   | 4680 |
| atro:         | aacccc | gtaacacac  | acttatcocc   | actggcagca              | gecactggta   | acaggattag   | 4740 |
| Cause<br>Dece | acasac | tatgtaggc  | qtqctacaqa   | gttcttgaac              | tggtggccta   | actacggcta   | 4800 |
| cact          | agaaga | acaghabbh  | gtatctqcq    | tetgetgaac              | ccagttacct   | tcggaaaaag   | 4860 |
| cacc          | ~>~~   |            | ,            |                         |              |              |      |

```
agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg 4920
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 4980
ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc 5040
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag 5100
tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 5160
aggatetqt ctatttegtt catecatagt tgeetgacte gggggggggg ggegetgagg 5220
tetgeetegt gaagaaggtg ttgetgaete ataccaggee tgaategeee cateatecag 5280 -
ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc agttggtgat 5340
cttcaactca gcaaaagttc gatttattca acaaagccgc cgtcccgtca agtcagcgta 5460
atgctctgcc agtgttacaa ccaattaacc aattctgatt agaaaaactc atcgagcatc 5520
aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5580
ttctqtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag atcctggtat 5640
cggtctgcga ttccgactcg tccaacatca atacaaccta ttaatttccc ctcgtcaaaa 5700
ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga gaatggcaaa 5760
agettatgea titettteea gaettgitea acaggecage cattaegete gicateaaaa 5820
teactequat caaccaaacc gttattcatt egtgattgeg cetgagegag acgaaatacg 5880
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact 5940
qccaqcqcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6000
gttttcccgg ggatcgcagt ggtgagtaac catgcatcat caggagtacg gataaaatgc 6060
ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat ctcatctgta 6120
acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc 6180
ccatacaatc gatagattgt cgcacctgat tgcccgacat tatcgcgagc ccatttatac 6240
ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga cgtttcccgt 6300
tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag ttttattgtt 6360
catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6420
ctttccccc cccccatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 6480
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 6540
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6600
atcacgaggc cctttcgtc
```

```
<210> 25
<211> 6655
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
```

## <400> 25

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 teggegggtg teggggetgg ettaactatg eggeateaga geagattgta etgagagtge 180 accatatgeg gtgtgaaata eegecaagat gegtaaggag aaaatacege ateagattgg 240 etattggeea ttgeateeat tgtateeata teatatatg tacatttata ttggeteatg 300 tecaacatta eegecatgtt gacattgatt attgaetagt tattaatagt aateaattae 360 ggggteatta gtteatagee eatatatgga gtteeeggtt accataatta eeggtaaatgg 420 eeegeetgge tgaeegeea aegaeeeeeg eeeattgaeg teaataatga egtatgtee 480 eatagtaacg eeatatagga ettteeattg aegteaatg gtggagtatt taeggtaaa 540 tgaeeggtaaa teggeeeeet ggeattatge eeagtaeatg geggatatt taggeeggt 180 aegteaatgg 240 eeggeggataa teggeeeeet ggeattatge eeagtaeatg gtggagtatt taeggtaaa 600 tgaeeggtaaa teggeeeet taggeatee tattaeeatg gtgatgeggt tettggeagta 720 eateaatggg egtggatage ggtttgaeet aeggggattt eeeagteee aeeeettga 780

| WU 02/3294. | )                        |            |            |            | PC1/US     | 01/25/2 |
|-------------|--------------------------|------------|------------|------------|------------|---------|
| cgtcaatggg  | agtttgtttt               | ggcaccaaaa | tcaacgggac | tttccaaaat | gtcgtaacaa | 840     |
|             | ttgacgcaaa               |            |            |            |            |         |
|             | gtgaaccgtc               |            |            |            |            |         |
|             | cgggaccgat               |            |            |            |            |         |
|             | aagagtgacg               |            |            |            |            |         |
|             | gctatactgt               |            |            |            |            |         |
|             | tatagcttag               |            |            |            |            |         |
|             | gatactttcc               |            |            |            |            |         |
|             | atgccaatac               |            |            |            |            |         |
|             | ccatttatta               |            |            |            |            |         |
|             | attaaacata               |            |            |            |            |         |
|             | tctccggtag               |            |            |            |            |         |
|             | ggtcgctcgg               |            |            |            |            |         |
|             | ccaccaccac               |            |            |            |            |         |
| gaaaatgagc  | gtggagattg               | ggctcgcacg | gctgacgcag | atggaagact | taaggcagcg | 1680    |
|             | atgcaggcag               |            |            |            |            |         |
|             | tgttaacggt               |            |            |            |            |         |
|             | gacataatag               |            |            |            |            |         |
|             | gtcgtcgaca               |            |            |            |            |         |
|             | ccagcacctg               |            |            |            |            |         |
| tgctgatgat  | ctgcagcgcc               | accgagaagc | tgtgggtgac | cgtgtactac | ggcgtgcccg | 2040    |
| tgtggaagga  | ggccaccacc               | accctgctct | gcgccagcga | cgccaaggcc | tacgacaccg | 2100    |
| aggtgcacaa  | cgtgtgggcc               | acccacgcct | gcgtgcccac | cgaccccaac | ccccaggagg | 2160    |
| tggtgctggt  | gaacgtgacc               | gagaacttcg | acatgtggaa | gaacgacatg | gtggagcaga | 2220    |
| tgcacgagga  | catcatcagc               | ctgtgggacc | agagcctgaa | gccctgcgtg | aagctgaccc | 2280    |
|             | gagcctgaag               |            |            |            |            |         |
|             | cggcaaggtg               |            |            |            |            |         |
|             | cgacaccacc               |            |            |            |            |         |
|             | caaggtgagc               |            |            |            |            |         |
|             | gaagtgcaag               |            |            |            |            |         |
|             | gtgcacccac               |            |            |            |            |         |
|             | cgaggaggag               |            |            |            |            |         |
|             | gcagctgaac               |            |            |            |            |         |
|             | ccgcgccaag               |            |            |            |            |         |
|             | caacaagacc               |            |            |            |            |         |
|             | cttcaactgc               |            |            |            |            |         |
|             | gttcaacagc               |            |            |            |            |         |
|             | cctgccctgc               |            |            |            |            |         |
|             | cgccccccc                |            |            |            |            |         |
|             | ccgcgacggc               |            |            |            |            |         |
| gcggcgacat  | gcgcgacaac               | tggcgcagcg | agetgtaeaa | gcacaaggtg | gtgaagatcg | 3240    |
|             | cgtggccccc               |            |            |            |            |         |
|             | gcagcagcag               |            |            |            |            |         |
|             | cgtgtggggc               |            |            |            |            |         |
|             | ccagcagcag               |            |            |            |            |         |
|             | caacaactac<br>gaacgagcag |            |            |            |            |         |
|             | caccaactgg               |            |            |            |            |         |
|             | tgtttgcccc               |            |            |            |            |         |
|             | ctaataaaat               |            |            |            |            |         |
|             | tggggtgggg               |            |            |            |            |         |
|             | tgcggtgggc               |            |            |            |            |         |
| cctagaccaa  | aaagaagcag               | gracaterer | ttetetatas | cacaccctot | ccacgcccct | 3960    |
|             | tccagcccca               |            |            |            |            |         |
| toccacoco   | taaagtactt               | agagggatch | ctccctccct | catcagecea | ccaaaccaa  | 4080    |
| Localoge    | Janageacee               | 224252256  |            | agecoa     |            |         |

```
cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagtg cagagggaga 4140
qaaaatqcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt aaggccatga 4200
tttaaggcca tcatggcctt aatcttccgc ttcctcgctc actgactcgc tgcgctcggt 4260
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 4320
atcaggggat aacgcaggaa agaacatgtg agcaaaaaggc cagcaaaagg ccaggaaccg 4380
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa 4440
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 4500
tecceetgga ageteceteg tgegetetee tgtteegace etgeegetta eeggataeet 4560
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 4620
cagtteggtg taggtegtte getecaaget gggetgtgtg cacgaacece eegtteagee 4680
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 4740
ategecactg geageageca etggtaacag gattageaga gegaggtatg taggeggtge 4800
tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 4860
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 4920
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 4980
aaaaggatet caagaagate etttgatett ttetaegggg tetgaegete agtggaaega 5040
aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 5100
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 5160
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 5220
catagttgcc tgactcgggg ggggggggcg ctgaggtctg cctcgtgaag aaggtgttgc 5280
tgactcatac caggootgaa togooccatc atocagocag aaagtgaggg agccacggtt 5340
gatgagaget ttgttgtagg tggaccagtt ggtgattttg aacttttget ttgccacgga 5400
acggtctgcg ttgtcgggaa gatgcgtgat ctgatccttc aactcagcaa aagttcgatt 5460
tattcaacaa agccgccgtc ccgtcaagtc agcgtaatgc tctgccagtg ttacaaccaa 5520
ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa tttattcata 5580
tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg agaaaactca 5640
ccgaggcagt tccataggat ggcaagatcc tggtatcggt ctgcgattcc gactcgtcca 5700
acatcaatac aacctattaa tttcccctcg tcaaaaataa ggttatcaag tgagaaatca 5760
ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc tttccagact 5820
tqttcaacaq qccaqccatt acgctcgtca tcaaaatcac tcgcatcaac caaaccgtta 5880
ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa aggacaatta 5940
caaacaggaa togaatgcaa coggogoagg aacactgcca gogoatcaac aatattttca 6000
cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat cgcagtggtg 6060
agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag aggcataaat 6120
tecgteagee agtitagtet gaccatetea tetgtaacat cattggcaac getacetttg 6180
ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata gattgtcgca 6240
ectqattqcc cgacattatc gcgagcccat ttatacccat ataaatcagc atccatgttg 6300
gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat aacacccctt 6360
gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt tttatcttgt 6420
gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc ccattattga 6480
agcatttatc agggttattg totcatgagc ggatacatat ttgaatgtat ttagaaaaat 6540
aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 6600
attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtc
```

```
<210> 26
<211> 6658
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 26
```

| tegegegttt | cggtgatgac | ggtgaaaacc  | tetgacacat | gcagctcccg  | gagacggtca | 60   |
|------------|------------|-------------|------------|-------------|------------|------|
| cagcttgtct | gtaagcggat | gccgggagca  | gacaagcccg | tcagggcgcg  | tcagcgggtg | 120  |
| ttggcgggtg | teggggetgg | cttaactatg  | cggcatcaga | gcagattgta  | ctgagagtgc | 180  |
| accatatgcg | gtgtgaaata | ccgcacagat  | gcgtaaggag | aaaataccgc  | atcagattgg | 240  |
| ctattggcca | ttgcatacgt | tgtatccata  | tcataatatg | tacatttata  | ttggctcatg | 300  |
| tccaacatta | ccgccatgtt | gacattgatt  | attgactagt | tattaatagt  | aatcaattac | 360  |
|            |            |             |            |             | cggtaaatgg |      |
|            |            |             |            |             | cgtatgttcc |      |
|            |            |             |            |             | tacggtaaac |      |
| tgcccacttg | gcagtacatc | aagtgtatca  | tatgccaagt | acgcccccta  | ttgacgtcaa | 600  |
| tgacggtaaa | tggcccgcct | ggcattatgc  | ccagtacatg | accttatggg  | actttcctac | 660  |
|            |            |             |            |             | tttggcagta |      |
|            |            |             |            |             | accccattga |      |
| cgtcaatggg | agtttgtttt | ggcaccaaaa  | tcaacgggac | tttccaaaat  | gtcgtaacaa | 840  |
|            |            |             |            |             | atataagcag |      |
|            |            |             |            |             | ttgacctcca |      |
| tagaagacac | cgggaccgat | ccagcctccg  | cggccgggaa | cggtgcattg  | gaacgcggat | 1020 |
| tccccgtgcc | aagagtgacg | taagtaccgc  | ctatagactc | tataggcaca  | cccctttggc | 1080 |
| tcttatgcat | gctatactgt | ttttggcttg  | gggcctatac | acccccgctt  | ccttatgcta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg  | tgggttattg | accattattg  | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc  | cataacatgg | ctctttgcca  | caactatctc | 1260 |
| tattggctat | atgccaatac | tetgteette  | agagactgac | acggactctg  | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt  | cacatataca | acaacgccgt  | ccccgtgcc  | 1380 |
| cgcagttttt | attaaacata | gcgtgggatc  | tccacgcgaa | tetegggtae  | gtgttccgga | 1440 |
| catgggctct | tctccggtag | cggcggagct  | tccacatccg | agccctqqtc  | ccatgcctcc | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg  | ctcctaacag | tggaggccag  | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg  | cacaaggccg | tggcggtagg  | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg  | gctgacgcag | atggaagact  | taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt  | gtattctgat | aaqaqtcaqa  | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt | ggagggcagt  | gtagtctgag | cagtactcgt  | tgctgccgcg | 1800 |
| cgcgccacca | gacataatag | ctgacagact  | aacagactgt | tcctttccat  | gggtcttttc | 1860 |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca  | gatatcgcgg | ccactctaga  | atgcgcgtga | 1920 |
| aggagaagta | ccagcacctg | tggcgctggg  | qctqqcqctq | gggcaccatg  | ctgctgggca | 1980 |
| tgctgatgat | ctgcagcgcc | accgagaagc  | tatagataac | cgtgtactac  | ggcgtgcccg | 2040 |
| tgtggaagga | ggccaccacc | accetquetet | qcqccaqcqa | cgccaaggcc  | tacgacaccg | 2100 |
| aggtgcacaa | cgtgtgggcc | acceacqcct  | gcgtgcccac | cgaccccaac  | ccccaggagg | 2160 |
| tggtgctggt | gaacgtgacc | gagaacttcg  | acatgtggaa | gaacgacatg  | gtggagcaga | 2220 |
| tgcacgagga | catcatcage | ctgtgggacc  | agageetgaa | accetacata  | aagctgaccc | 2280 |
| ccctgtgcgt | gagectgaag | tgcaccgacg  | ctaqcaaqaa | ctgcagcttc  | aacatcagca | 2340 |
| ccaqcatccq | cagcaagata | cagaaggagt  | acgcettett | ctacaagetg  | gacatcatcc | 2400 |
| ccatcgacaa | cgacaccacc | agctacagcc  | tgaccagetg | caacaccagc  | gtgatcaccc | 2460 |
| aggcctgccc | caaggtgagc | ttcgagccca  | tccccaacca | ctactgcgcc  | cccgccggct | 2520 |
| togocatoct | gaagtgcaag | gacaagaagt  | tcaacggcaa | agacccctac  | accaacgtga | 2580 |
| gcaccgtgca | gtgcacccac | ggcatccqcc  | ccataataaa | cacccagctg  | ctgctgaacg | 2640 |
| gtagectgge | cqaqqaqqaq | gtggtgatcc  | gcagcgctaa | cttcaccaac  | aacgccaagg | 2700 |
| tgatcatcgt | gcagctgaac | gagagcatag  | agatcaactg | cacccaccc   | aacaacaaca | 2760 |
| cccqcaaqaq | catccacatc | aaccccaacc  | acacetteta | caccaccode  | gagatcatcg | 2820 |
| gcgacatccg | ccaggcccac | tgcaacctga  | accacaccaa | atagaacaac  | accetgaaca | 2880 |
| agatcataat | caagetgege | gagcagttcg  | gcaacaagac | catcototto  | aagcacagca | 2940 |
| acadcaacaa | ccccgagate | gtgacccaca  | gcttcaacta | caacaacasa  | ttcttctact | 3000 |
| gcaacagege | tageetgeen | taccacatea  | agcagatrat | caacatotoo  | cagaaggtgg | 3060 |
| gcaaggcat  | ghacgecee  | cccatcacco  | accadatece | ctocaccacc  | aacatcaccg | 3120 |
| acctactact | gacccgcgad | adcadcaaca  | gcaacaacca | gagcgagcagc | ttccgtccgg | 3150 |
| deddeddea  | catgcgcgac | aactggcgca  | gcgagctgta | caagtacaag  | gtggtgaaga | 3240 |
| tegagecect | dadcataacc | cccaccaaoo  | ccaagettae | catccadag   | cgccagctgc | 3230 |
|            | 222-2-3300 |             |            | -geraggee   | caccagety  | 2200 |

| tgagcggcat | cgtgcagcag | cagaacaacc | tgctgcgcgc | catcgaggcc  | cagcagcacc | 3360 |
|------------|------------|------------|------------|-------------|------------|------|
| tgctgcagct | gaccgtgtgg | ggcatcaagc | agctgcaggc | ccgcaccctg  | gccgtggagc | 3420 |
| gctacctgaa | ggaccagcag | cagctcgagc | agatctggaa | ccacaccacc  | tggatggagt | 3480 |
| gggaccgcga | gatcaacaac | tacaccagcc | tgatccacag | cctgatcgag  | gagagccaga | 3540 |
| accagcagga | gaagaacgag | caggagetge | tggagctgga | caagtgggcc  | agcctgtgga | 3600 |
| actggttcaa | catcaccaac | tggctgtggt | gaggatccag | atctgctgtg  | ccttctagtt | 3660 |
| gccagccatc | tgttgtttgc | ccctcccccg | tgccttcctt | gaccctggaa  | ggtgccactc | 3720 |
| ccactgtcct | ttcctaataa | aatgaggaaa | ttgcatcgca | ttgtctgagt  | aggtgtcatt | 3780 |
| ctattctggg | gggtggggtg | gggcaggaca | gcaaggggga | ggattgggaa  | gacaatagca | 3840 |
| ggcatgctgg | ggatgcggtg | ggctctatgg | gtacccaggt | gctgaagaat  | tgacccggtt | 3900 |
| cctcctgggc | cagaaagaag | caggcacatc | cccttctctg | 'tgacacaccc | tgtccacgcc | 3960 |
| cctggttctt | agttccagcc | ccactcatag | gacactcata | gctcaggagg  | gctccgcctt | 4020 |
| caatcccacc | cgctaaagta | cttggagcgg | tctctcctc  | cctcatcagc  | ccaccaaacc | 4080 |
| aaacctagcc | tccaagagtg | ggaagaaatt | aaagcaagat | aggctattaa  | gtgcagaggg | 4140 |
| agagaaaatg | cctccaacat | gtgaggaagt | aatgagagaa | atcatagaat  | tttaaggcca | 4200 |
| tgatttaagg | ccatcatggc | cttaatcttc | cgcttcctcg | ctcactgact  | cgctgcgctc | 4260 |
| ggtcgttcgg | ctgcggcgag | cggtatcagc | tcactcaaag | gcggtaatac  | ggttatccac | 4320 |
| agaatcaggg | gataacgcag | gaaagaacat | gtgagcaaaa | ggccagcaaa  | aggccaggaa | 4380 |
|            |            |            |            |             | acgagcatca |      |
| caaaaatcga | cgctcaagtc | agaggtggcg | aaacccgaca | ggactataaa  | gataccaggc | 4500 |
| gtttcccct  | ggaagctccc | tegtgegete | tcctgttccg | accctgccgc  | ttaccggata | 4560 |
| cctgtccgcc | tttctccctt | cgggaagcgt | ggcgctttct | catageteae  | gctgtaggta | 4620 |
| tctcagttcg | gtgtaggtcg | ttcgctccaa | gctgggctgt | gtgcacgaac  | cccccgttca | 4680 |
| gcccgaccgc | tgcgccttat | ccggtaacta | tcgtcttgag | tccaacccgg  | taagacacga | 4740 |
| cttatcgcca | ctggcagcag | ccactggtaa | caggattagc | agagcgaggt  | atgtaggcgg | 4800 |
| tgctacagag | ttcttgaagt | ggtggcctaa | ctacggctac | actagaagaa  | cagtatttgg | 4860 |
| tatctgcgct | ctgctgaagc | cagttacctt | cggaaaaaga | gttggtagct  | cttgatccgg | 4920 |
| caaacaaacc | accgctggta | gcggtggttt | ttttgtttgc | àagcagcaga  | ttacgcgcag | 4980 |
| aaaaaaagga | tctcaagaag | atcctttgat | cttttctacg | gggtctgacg  | ctcagtggaa | 5040 |
| cgaaaactca | cgttaaggga | ttttggtcat | gagattatca | aaaaggatct  | tcacctagat | 5100 |
| ccttttaaat | taaaaatgaa | gttttaaatc | aatctaaagt | atatatgagt  | aaacttggtc | 5160 |
| tgacagttac | caatgcttaa | tcagtgaggc | acctatctca | gcgatctgtc  | tatttcgttc | 5220 |
| atccatagtt | gcctgactcg | 999999999  | gcgctgaggt | ctgcctcgtg  | aagaaggtgt | 5280 |
| tgctgactca | taccaggcct | gaatcgcccc | atcatccagc | cagaaagtga  | gggagccacg | 5340 |
| gttgatgaga | gctttgttgt | aggtggacca | gttggtgatt | ttgaactttt  | gctttgccac | 5400 |
| ggaacggtct | gcgttgtcgg | gaagatgcgt | gatctgatcc | ttcaactcag  | caaaagttcg | 5460 |
| atttattcaa | caaagccgcc | gtcccgtcaa | gtcagcgtaa | tgctctgcca  | gtgttacaac | 5520 |
| caattaacca | attctgatta | gaaaaactca | tcgagcatca | aatgaaactg  | caatttattc | 5580 |
| atatcaggat | tatcaatacc | atatttttga | aaaagccgtt | tctgtaatga  | aggagaaaac | 5640 |
| tcaccgaggc | agttccatag | gatggcaaga | tcctggtatc | ggtctgcgat  | tccgactcgt | 5700 |
| ccaacatcaa | tacaacctat | taatttcccc | tcgtcaaaaa | taaggttatc  | aagtgagaaa | 5760 |
| tcaccatgag | tgacgactga | atccggtgag | aatggcaaaa | gcttatgcat  | ttctttccag | 5820 |
| acttgttcaa | caggccagcc | attacgctcg | tcatcaaaat | cactcgcatc  | aaccaaaccg | 5880 |
| ttattcattc | gtgattgcgc | ctgagcgaga | cgaaatacgc | gatcgctgtt  | aaaaggacaa | 5940 |
| ttacaaacag | gaatcgaatg | caaccggcgc | aggaacactg | ccagcgcatc  | aacaatattt | 6000 |
| tcacctgaat | caggatattc | ttctaatacc | tggaatgctg | ttttcccggg  | gatcgcagtg | 6060 |
| gtgagtaacc | atgcatcatc | aggagtacgg | ataaaatgct | tgatggtcgg  | aagaggcata | 6120 |
| aattccgtca | gccagtttag | tctgaccatc | tcatctgtaa | catcattggc  | aacgctacct | 6180 |
| ttgccatgtt | tcagaaacaa | ctctggcgca | tegggettee | catacaatcg  | atagattgtc | 6240 |
| gcacctgatt | gcccgacatt | atcgcgagcc | catttatacc | catataaatc  | agcatccatg | 6300 |
| ttggaattta | atcgcggcct | cgagcaagac | gtttcccgtt | gaatatggct  | cataacaccc | 6360 |
| cttgtattac | tgtttatgta | agcagacagt | tttattgttc | atgatgatat  | atttttatct | 6420 |
| tgtgcaatgt | aacatcagag | attttgagac | acaacgtggc | tttcccccc   | ccccattat  | 6480 |
| tgaagcattt | atcagggtta | ttgtctcatg | agcggataca | catttgaatg  | tatttagaaa | 6540 |
| aataaacaaa | taggggttcc | gcgcacattt | ccccgaaaag | tgccacctga  | cgtctaagaa | 6600 |
|            |            |            |            |             |            |      |

accattatta tcatqacatt aacctataaa aataqqcqta tcacqagqcc ctttcqtc 6658

<210> 27 <211> 6616 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tegegegttt eggtgatgae ggtgaaaacc tetgacacat geageteeeg gagaeggtea 60 cagettgtct gtaageggat geegggagea gaeaageeeg teagegggeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccqcctqqc tqaccqccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480. catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 aqctcqttta qtqaaccqtc aqatcqcctg gagacqccat ccacqctgtt ttgacctcca 960 taqaaqacac cqqqaccqat ccaqcctccq cqqccgggaa cggtgcattg gaacgcggat 1020 tecceqtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactqt ttttggettg gggeetatac acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920 aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040 tgtggaagga ggccaccacc accetgetet gegccagcga cgccaaggcc tacgacaccg 2100 aggtgcacaa cgtgtgggcc acccacgcct gcgtgcccac cgaccccaac ccccaggagg 2160 tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220 tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280 ccctqtgcqt gagcctqaaq tqcaccgacc tgaagaacga caccaacacc aacagcagca 2340 qcggccgcat gatcatqqag aagggcgaga tcaagaactg cagcttcgct agcaccagct 2400 quaacaccag cgtqatcacc caggcctgcc ccaaggtgag cttcgagccc atccccaacc 2460 actactgcgc ccccgccggc ttcgccatcc tgaagtgcaa ggacaagaag ttcaacggca 2520

```
agggeceetg caccaacgtg agcacegtge agtgeaceca eggeateege eeegtggtga 2580
gcacccaget getgetgaac ggtageetgg eegaggagga ggtggtgate egeagegeta 2640
acttegeega caaegeeaag gtgateateg tgeagetgaa egagagegtg gagateaact 2700
gcacccgcgc tagcgcccac tgcaacctga gccgcgccaa gtggaacgac accctgaaca 2760
agatcgtgat caagctgcgc gagcagttcg gcaacaagac catcgtgttc aagcacagca 2820
geggeggega ccccgagate gtgacccaca getteaactg eggeggegag ttettetact 2880
gcaacagcac ccagctgttc aacagcacct ggttcaacag cacctggagc accgagggca 2940
acatgtggca gaaggtgggc aaggccatgt acgcccccc catcagcggc cagatccgct 3060
qcaqcaqcaa catcaccggc ctgctgctga cccgcgacgg cggcaacagc aacaacgaga 3120
gcgagatett ccgtccgggc ggcggcgaca tgcgcgacaa ctggcgcagc gagctgtaca 3180
agtacaaggt ggtgaagatc gagcccctgg gcgtggcccc caccaaggcc aagcttaccg 3240
tocaggodog coagotgotg agoggoatog tgoagoagoa gaacaacotg otgogogoda 3300
tegaggeeca geageacetg etgeagetga eegtgtgggg cateaageag etgeaggeec 3360
qcaccetqqc cgtggagcgc tacetgaagg accageagca getegagcag atetggaace 3420
acaccacctg gatggagtgg gaccgcgaga tcaacaacta caccagcctg atccacagcc 3480
tgatcgagga gagccagaac cagcaggaga agaacgagca ggagctgctg gagctggaca 3540
aqtqqqccaq cctqtggaac tggttcaaca tcaccaactg gctgtggtga ggatccagat 3600
ctgctgtgcc ttctagttgc cagccatctg ttgtttgccc ctccccgtg ccttccttga 3660
ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt gcatcgcatt 3720
attgggaaga caatagcagg catgctgggg atgcggtggg ctctatgggt acccaggtgc 3840
tgaagaattg acceggttcc tcctgggcca gaaagaagca ggcacatccc cttctctgtg 3900
acacaccety tecacycece tygttettay ttecayeece acteatagya cacteatage 3960
traggaggge tregerttea attreacting chaaagtact tggageggte trtecettee 4020
tcatcaqccc accaaaccaa acctagcctc caagagtggg aagaaattaa agcaagatag 4080
gctattaagt gcagagggag agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat 4140
catagaattt taaggccatg atttaaggcc atcatggcct taatcttccg cttcctcgct 4200
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagetc actcaaaggc 4260
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 4320
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 4380
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 4440
actataaaga taccaggogt ttccccctgg aagctccctc gtgcgctctc ctgttccgac 4500
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 4560
tageteaege tgtaggtate teagtteggt gtaggtegtt egeteeaage tgggetgtgt 4620
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 4680
caacceggta agacacgact tategceact ggcagcagce actggtaaca ggattagcag 4740
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 4800
tagaagaaca gtatttggta tetgegetet getgaageea gttacetteg gaaaaagagt 4860
tggtagetet tgateeggea aacaaaceae egetggtage ggtggttttt ttgtttgeaa 4920
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 4980
qtctqacqct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 5040
aaggatette acetagatee tittaaatta aaaatgaagt titaaateaa tetaaagtat 5100
atatgaqtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 5160
gatctgtcta tttcgttcat ccatagttgc ctgactcggg gggggggggc gctgaggtct 5220
gcctcgtgaa gaaggtgttg ctgactcata ccaggcctga atcgccccat catccagcca 5280
gaaagtgagg gagccacggt tgatgagagc tttgttgtag gtggaccagt tggtgatttt 5340
gaacttttgc tttgccacgg aacggtctgc gttgtcggga agatgcgtga tctgatcctt 5400
caactcagca aaagttegat ttattcaaca aageegeegt ceegtcaagt cagegtaatg 5460
ctctgccagt gttacaacca attaaccaat tctgattaga aaaactcatc gagcatcaaa 5520
tqaaactqca atttattcat atcaggatta tcaataccat atttttgaaa aagccgtttc 5580
tgtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg 5640
totgogatto ogactogtoc aacatoaata caacotatta atttoccoto gtoaaaaata 5700
aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc 5760
ttatqcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca 5820
```

```
ctcgcatcaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg aaatacgcga 5880 tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgcag gaacactgcc 5940 agcgcatcaa caatatttc acctgaatca ggatattctt ctaatacctg gaatgctgtt 6000 ttcccgggga tcgcagtggt gagtaaccat gcatcatcag gagtacggat aaaatgcttg 6060 atggtcggaa gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca 6120 tcaatggcaa cgctacctt gccatgttc agaacaact ctggcgcatc gggcttccca 6180 tacaatcgat agattgtcg acctgattgc ccgacattat cgcggagccca tttataccca 6240 tataaatcag catccatgtt ggaatttaat cgcggcctcg agcaagacgt ttcccgttga 6300 atatggctca taacacccct tgtattactg tttatgtaag cagcaagacgt ttcccgttga 6300 atatggctca ttttatcttg tgcaatgtaa catcagggat tttgagacac aacgtggctt 6420 tcccccccc cccattattg aagcattat cagggttatt gtctcatgag cggatacata 6480 tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 6540 acgagggccct ttcgtc
```

```
<210> 28
<211> 6619
<212> DNA
<213> Artificial Sequence
<223> plasmid pVR1012x/s containing HIV genes
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960
taqaaqacac cqqqaccqat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tottatgcat getatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattqqctat atqccaatac tctqtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cqcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560
agcacaatge ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
qcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
```

|            | •                    |              |              |            |                          |      |
|------------|----------------------|--------------|--------------|------------|--------------------------|------|
| gttgcggtgc | tgttaacggt           | ggagggcagt   | gtagtctgag   | cagtactcgt | tgctgccgcg               | 1800 |
| cgcgccacca | gacataatag           | ctgacagact   | aacagactgt   | tectttecat | gggtcttttc               | 1860 |
| tgcagtcacc | gtcgtcgaca           | cgtgtgatca   | gatatcgcgg   | ccgctctaga | atgcgcgtga               | 1920 |
| aggagaagta | ccagcacctg           | tggcgctggg   | gctggcgctg   | gggcaccatg | ctgctgggca               | 1980 |
| tgctgatgat | ctgcagcgcc           | accgagaagc   | tgtgggtgac   | cgtgtactac | ggcgtgcccg               | 2040 |
| tqtqqaaqqa | ggccaccacc           | accctgctct   | gcgccagcga   | cgccaaggcc | tacgacaccg               | 2100 |
| aggtgcacaa | cgtgtgggcc           | acccacgcct   | gcgtgcccac   | cgaccccaac | ccccaggagg               | 2160 |
| taatactaat | qaacqtqacc           | gagaacttcg   | acatgtggaa   | gaacgacatg | gtggagcaga               | 2220 |
| tocacgagga | catcatcage           | ctgtgggacc   | agagcctgaa   | gccctgcgtg | aagctgaccc               | 2280 |
| ccctatacat | gageetgaag           | tgcaccgacc   | tgaagaacga   | caccaacacc | aacagcagca               | 2340 |
| gcggcgcat  | gatcatggag           | aaqqqcqaqa   | tcaagaactg   | cagcttcgct | agcaccagct               | 2400 |
| gcaacaccag | cataatcacc           | caggeetgee   | ccaaggtgag   | cttcgagccc | atccccaacc               | 2460 |
| actactococ | ccccaccaac           | ttcgccatcc   | tgaagtgcaa   | ggacaagaag | ttcaacggca               | 2520 |
| agggccctg  | caccaacgtg           | agcaccqtqc   | agtgcaccca   | cggcatccgc | cccgtggtga               | 2580 |
| gcacccagct | actactaaac           | gatageetga   | ccqaqqaqqa   | ggtggtgatc | cgcagcgcta               | 2640 |
| acttcaccas | caacaccaaa           | gtgatcatcg   | tacaactaaa   | cgagagcgtg | gagatcaact               | 2700 |
| acecegecga | caacaacaac           | acccccaaga   | gcatccacat   | caaccccaac | cgcgccttct               | 2760 |
| agaggaggg  | caacatcatc           | ggcgacatcc   | gccaggccca   | ctgcaacctg | agccgcgcca               | 2820 |
| acaccaccgg | caccetgaac           | aagatcgtga   | tcaagetgcg   | cgagcagttc | ggcaacaaga               | 2880 |
| agiggaacga | caccegaac            | addaccaca    | accccaagat   | cataacccac | agcttcaact               | 2940 |
| ccaccgcgcc | attattataa           | tacaacaaca   | chacctacc    | chaccacate | aagcagatca               | 3000 |
| geggeggega | geeceecae            | ageaacageg   | tatacacccc   | cccatcage  | ggccagatcc               | 3060 |
| ccaacatgtg | gcagaaggcg           | ggcaaggcca   | taaccacaa    | caacaacaac | agcaacaacg               | 3120 |
| getgeageag | caacaccacc           | ggcccgccgc   | acatococoa   | caactggcgc | agcgagctgt               | 3180 |
| agagcgagac | ecteegteeg           | ggcggcggcg   | tagagataga   | ccccaccaad | accaaactta               | 3240 |
| acaagtacaa | ggcggcgaag           | ategageeee   | tagtagaga    | ccccaccaag | gccaagctta               | 3300 |
| ccgtccaggc | cegecagetg           | ctgageggea   | tanagatata   | gcagaacaac | ctgctgcgcg               | 3360 |
| ccatcgaggc | ccagcagcac           | ctgctgcagc   | agaccacaca   | gggcaccaag | cagctgcagg<br>cagatctgga | 3420 |
| cccgcaccct | ggeegtggag           | tagaaaaaa    | aggaccagca   | ctacaccac  | ctgatccaca               | 3480 |
| accacaccac | etggatggag           | rgggaccgcg   | agaccaacaa   | acaggageta | ctggagctgg               | 3540 |
| geetgatega | ggagagecag           | aactagtagg   | agauguucgu   | ctaactataa | tgaggatcca               | 3600 |
| acaagtgggc | cageetgegg           | taccaaccat   | ctattattta   | ccctcccc   | gtgccttcct               | 3660 |
| gatetgetgt | geettetagt           | egecagecae   | tttcctaata   | aaatgaggaa | attgcatcgc               | 3720 |
| tgaccctgga | aggrgeeact           | behaltetee   | aggataggat   | adacaaaaa  | agcaaggggg               | 3780 |
| attgtctgag | taggigical           | cocaccocgg   | ggggcggggc   | ggggcaggac | ageacogg                 | 3840 |
| aggattggga | agacaatage           | aggeargerg   | gggatgeggt   | gggcccaca  | ggtacccagg               | 3900 |
| tgctgaagaa | ttgacccggt           |              | teathanaa    | gcaggcacac | ccccttctct               | 3960 |
| gtgacacacc | ctgtccacgc           | ceetggttet   | ragiticage   | acttaccaca | ggacactcat               | 4020 |
| agctcaggag | ggctccgcct           | tcaatcccac   | cegetaaagt   | actiggageg | gtctctccct               | 4080 |
| ccctcatcag | cccaccaaac           | caaacccagc   | ccccaagage   | tatasaasaa | taaagcaaga               | 4140 |
| taggctatta | agtgcagagg           | gagagaaaat   | geetecaaca   | cattastatt | taatgagaga               | 4200 |
| aatcatagaa | ttttaaggee           | acgatttaag   | gecateatgg   | geggtatea  | ccgcttcctc               | 4260 |
| gctcactgac | tegetgeget           | cggccgccg    | gergeggega   | geggeaceag | ctcactcaaa               | 4320 |
| ggcggtaata | cggttatcca           | cagaaccagg   | ggataacgca   | gyaaayaaca | tgtgagcaaa               | 4380 |
| aggccagcaa | aaggccagga           | accgtaaaaa   | ggccgcgttg   | ceggegeeee | tccataggct               | 4300 |
| cegececect | gacgagcatc           | acaaaaatcg   | acgctcaagt   | cagaggegge | gaaacccgac               | 3330 |
| aggactataa | agataccagg           | gtttcccc     | tggaagetee   | ctcgtgcgct | ctcctgttcc               | 4500 |
| gaccctgccg | cttaccggat           | acctgtccgc   | ctttctccct   | ccgggaagcg | tggcgctttc               | 4500 |
| tcatagctca | cgctgtaggt           | atctcagtto   | ggtgtaggtc   | gttegeteea | agctgggctg               | 4620 |
| tgtgcacgaa | ceceegtte            | : agcccgaccg | ctgcgcctta   | cccggtaact | : atcgtcttga             | 4740 |
| gtccaacccg | gtaagacacg           | , acttategee | actggcagca   | gccactggta | acaggattag               | 4000 |
| cagagcgagg | tatgtaggcg           | , gtgctacaga | gttcttgaag   | rggrggccta | actacggcta               | 4800 |
| cactagaaga | acagtatttg           | gtatctgcgc   | : tctgctgaag | ccagttacct | tcggaaaaag               | 4850 |
| agttggtago | : tcttgatcc <u>c</u> | g gcaaacaaac | : caccgctggt | agcggtggt  | tttttgtttg               | 4920 |
| caagcagcag | , attacgcgca         | a gaaaaaaagg | , atctcaagaa | gatcetttga | tcttttctac               | 4380 |
| ggggtctgac | gctcagtgga           | a acgaaaacto | : acgttaaggg | attttggtca | tgagattatc               | 5040 |

```
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag 5100
tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 5160
tetgeetegt gaagaaggtg ttgetgaete ataccaggee tgaategeee catcatecag 5280
ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc agttggtgat 5340
cttcaactca gcaaaagttc gatttattca acaaagccgc cgtcccgtca agtcagcgta 5460
atgetetgee agtgttacaa ecaattaace aattetgatt agaaaaacte ategageate 5520
aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5580
ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag atcctggtat 5640
eggtetgega tteegacteg tecaacatea atacaaceta ttaattteee etegteaaaa 5700
ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga gaatggcaaa 5760
agettatgea tttettteca gaettgttea acaggecage cattacgete gteatcaaaa 5820
tcactcgcat caaccaaacc gttattcatt cgtgattgcg cctgagcgag acgaaatacg 5880
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact 5940
gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6000
gttttcccgg ggatcgcagt ggtgagtaac catgcatcat caggagtacg gataaaatgc 6060
ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat ctcatctgta 6120
acateattgg caacgetace thtgccatgt theagaaaca actetggege alegggette 6180
ccatacaatc gatagattgt cgcacctgat tgcccgacat tatcgcgagc ccatttatac 6240
ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga cgtttcccgt 6300
tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag ttttattgtt 6360
catgatgata tatttttatc ttqtqcaatg taacatcaga gattttqaqa cacaacqtqq 6420
ctttccccc cccccatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 6480
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 6540
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6600
atcacgaggc cctttcgtc
```

```
<210> 29
<211> 6634
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 29
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat ge
```

tegegegtht eggtgatgae ggtgaaaace tetgaeacat geageteeeg gagaeggtea 60 cagcttgtct gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cqtatqttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattqa 780 cgtcaatggg agtttgttt ggcaccaaaa tcaacgggac tttccaaaat qtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataaqcaq 900 agctegttta gtgaaccgtc agategeetg gagaegecat ceaegetgtt ttgaecteca 960

| tagaagacac | cgggaccgat | ccagcctccg | cggccgggaa | cggtgcattg | gaacgcggat | 1020 |
|------------|------------|------------|------------|------------|------------|------|
| tccccgtgcc | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca | cccctttggc | 1080 |
| tcttatgcat | gctatactgt | ttttggcttg | gggcctatac | acccccgctt | ccttatgcta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg | tgggttattg | accattattg | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260 |
| tattggctat | atgccaatac | tctgtccttc | agagactgac | acggactctg | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacgccgt | ccccgtgcc  | 1380 |
| cgcagtttt  | attaaacata | gcgtgggatc | tccacgcgaa | tctcgggtac | gtgttccgga | 1440 |
| catgggctct | tctccggtag | cggcggagct | tccacatccg | agccctggtc | ccatgcctcc | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact | taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg | 1800 |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | tcctttccat | gggtcttttc | 1860 |
|            | gtcgtcgaca |            |            |            |            |      |
|            | ccagcacctg |            |            |            |            |      |
|            | ctgcagcgcc |            |            |            |            |      |
|            | ggccaccacc |            |            |            |            |      |
|            | cgtgtgggcc |            |            |            |            |      |
|            | gaacgtgacc |            |            |            |            |      |
|            | catcatcagc |            |            |            |            |      |
|            | gagcctgaag |            |            |            |            |      |
|            | gatcatggag |            |            |            |            |      |
|            | caaggtgcag |            |            |            |            |      |
|            | caccaccagc |            |            |            |            |      |
|            | ggtgagcttc |            |            |            |            |      |
|            | gtgcaaggac |            |            |            |            |      |
|            | cacccacggc |            |            |            |            |      |
|            | ggaggaggtg |            |            |            |            |      |
|            | gctgaacgag |            |            |            |            |      |
|            | cgccaagtgg |            |            |            |            |      |
|            | caagaccatc |            |            |            |            |      |
|            | caactgcggc |            |            |            |            |      |
|            | gatcatcaac |            |            |            |            |      |
|            | gatccgctgc |            |            |            |            |      |
|            | caacgagagc |            |            |            |            |      |
| ggcgcagcga | gctgtacaag | tacaaggtgg | tgaagatcga | gcccctgggc | gtggccccca | 3240 |
| ccaaggccaa | gcttaccgtc | caggcccgcc | agctgctgag | cggcatcgtg | cagcagcaga | 3300 |
| acaacctgct | gcgcgccatc | gaggcccagc | agcacctgct | gcagctgacc | gtgtggggca | 3360 |
| tcaagcagct | gcaggcccgc | accctggccg | tggagcgcta | cctgaaggac | cagcagcagc | 3420 |
| tcgagcagat | ctggaaccac | accacctgga | tggagtggga | ccgcgagatc | aacaactaca | 3480 |
| ccagcctgat | ccacagcctg | atcgaggaga | gccagaacca | gcaggagaag | aacgagcagg | 3540 |
| agctgctgga | gctggacaag | tgggccagcc | tgtggaactg | gttcaacatc | accaactggc | 3600 |
| tgtggtgagg | atccagatct | gctgtgcctt | ctagttgcca | gccatctgtt | gtttgcccct | 3660 |
| ccccatacc  | ttccttgacc | ctggaaggtg | ccactcccac | tgtcctttcc | taataaaatg | 3720 |
| aggaaattgc | atcgcattgt | ctgagtaggt | gtcattctat | tctggggggt | ggggtggggc | 3780 |
|            | gggggaggat |            |            |            |            |      |
|            | ccaggtgctg |            |            |            |            |      |
|            | tctctgtgac |            |            |            |            |      |
|            | ctcatagete |            |            |            |            |      |
|            | tecetecete |            |            |            |            |      |
|            | caagataggc |            |            |            |            |      |
|            | agagaaatca |            |            |            |            |      |
| atcttccgct | tectegetea | ctgactcgct | gcgctcggtc | gttcggctgc | ggcgagcggt | 4260 |

```
atcageteae teaaaggegg taataeggtt atceaeagaa teaggggata aegeaggaaa 4320
gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc 4380
gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag 4440
gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt 4500
qeqetetect qttccqaccc tqccqcttac cqqatacctq tecqcctttc tcccttcggg 4560
aagegtggeg ettteteata geteaegetg taggtatete agtteggtgt aggtegtteg 4620
ctccaagetg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg 4680
taactategt ettgagteca acceggtaag acaegaetta tegecaetgg cageagecae 4740
tqqtaacaqq attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg 4800
gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt 4860
taccttegga aaaagagttg gtagetettg atceggeaaa caaaccaeeg etggtagegg 4920
tqqttttttt qtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 4980
tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt 5040
taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag 5160
tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct gactcggggg 5220
qqqqqqqqc tqaggtctgc ctcgtgaaga aggtgttgct gactcatacc aggcctgaat 5280
egececatea tecagecaga aagtgaggga gecaeggttg atgagagett tgttgtaggt 5340
qqaccaqttq qtgattttga acttttgctt tgccacggaa cggtctgcgt tgtcgggaag 5400
atgcqtqatc tqatccttca actcaqcaaa aqttcqattt attcaacaaa gccgccgtcc 5460
cqtcaaqtca qcqtaatqct ctqccaqtqt tacaaccaat taaccaattc tqattagaaa 5520
aactcatcqa qcatcaaatg aaactgcaat ttattcatat caggattatc aataccatat 5580
ttttgaaaaa gccqtttctq taatqaagqa gaaaactcac cgaggcagtt ccataggatg 5640
qcaaqatect ggtateggte tgegatteeg actegteeaa cateaataca acetattaat 5700
ttcccctcgt caaaaataag gttatcaagt gagaaatcac catgagtgac gactgaatcc 5760
ggtgagaatg gcaaaagctt atgcatttct ttccagactt gttcaacagg ccagccatta 5820
egetegteat caaaateact egeateaace aaacegttat teattegtga ttgegeetga 5880
gcqaqacgaa atacgcgatc gctgttaaaa ggacaattac aaacaggaat cgaatgcaac 5940
eggegeagga acactgecag egcateaaca atatttteac etgaateagg atattettet 6000
aatacctgga atgctgtttt cccggggatc gcagtggtga gtaaccatgc atcatcagga 6060
gtacggataa aatgcttgat ggtcggaaga ggcataaatt ccgtcagcca gtttagtctg 6120
accatctcat ctgtaacatc attggcaacg ctacctttgc catgtttcag aaacaactct 6180
ggcgcatcgg gcttcccata caatcgatag attgtcgcac ctgattgccc gacattatcg 6240
cgagcccatt tatacccata taaatcagca tccatgttgg aatttaatcg cggcctcgag 6300
caagacgttt cccgttgaat atggctcata acaccccttg tattactgtt tatgtaagca 6360
gacagtttta ttgttcatga tgatatattt ttatcttgtg caatgtaaca tcagagattt 6420
tgagacacaa cgtggctttc ccccccccc cattattgaa gcatttatca gggttattgt 6480
ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc 6540
acatttcccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat gacattaacc 6600
tataaaaata ggcgtatcac gaggcccttt cgtc
                                                                 6634
```

```
<210> 30
<211> 6538
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 30
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
```

|            |            |            |            |            | atcagattgg               |      |
|------------|------------|------------|------------|------------|--------------------------|------|
|            |            |            |            |            | ttggctcatg               |      |
|            |            |            |            |            | aatcaattac               |      |
|            |            |            |            |            | cggtaaatgg               |      |
| cccgcctggc | tgaccgccca | acgacccccg | cccattgacg | tcaataatga | cgtatgttcc               | 480  |
| catagtaacg | ccaataggga | ctttccattg | acgtcaatgg | gtggagtatt | tacggtaaac               | 540  |
| tgcccacttg | gcagtacatc | aagtgtatca | tatgccaagt | acgcccccta | ttgacgtcaa               | 600  |
|            |            |            |            |            | actttcctac               |      |
|            |            |            |            |            | tttggcagta               |      |
|            |            |            |            |            | accccattga               |      |
|            |            |            |            |            | gtcgtaacaa               |      |
|            |            |            |            |            | atataagcag               |      |
|            |            |            |            |            | ttgacctcca               |      |
|            |            |            |            |            | gaacgcggat               |      |
|            |            |            |            |            | ccctttggc                |      |
|            |            |            |            |            | ccttatgcta               |      |
|            |            |            |            |            | accactcccc               |      |
|            |            |            |            |            | caactatctc               |      |
|            |            |            |            |            | tatttttaca               |      |
|            |            |            |            |            | ccccgtgcc                |      |
|            |            |            |            |            | gtgttccgga               |      |
|            |            |            |            |            | ccatgcctcc               |      |
|            |            |            |            |            |                          |      |
|            |            |            |            |            | acttaggcac               |      |
|            |            |            |            |            | gtatgtgtct<br>taaggcagcg |      |
|            |            |            |            |            |                          |      |
|            |            |            |            |            | ggtaactccc               |      |
|            |            |            |            |            | tgctgccgcg               |      |
|            |            |            |            |            | gggtcttttc               |      |
|            |            |            |            |            | atgcgcgtga               |      |
|            |            |            |            |            | ctgctgggca               |      |
|            |            |            |            |            | ggegtgeeeg               |      |
|            |            |            |            |            | tacgacaccg               |      |
|            |            |            |            |            | ccccaggagg               |      |
|            |            |            |            |            | gtggagcaga               |      |
|            |            |            |            |            | aagctgaccc               |      |
|            |            |            |            |            | agcgtgatca               |      |
|            |            |            |            |            | gcccccgccg               |      |
|            |            |            |            |            | tgcaccaacg               |      |
|            |            |            |            |            | ctgctgctga               |      |
|            |            |            |            |            | gacaacgcca               |      |
|            |            |            |            |            | gctagcgccc               |      |
|            |            |            |            |            | atcaagctgc               |      |
|            |            |            |            |            | gaccccgaga               |      |
|            |            |            |            |            | acccagctgt               |      |
|            |            |            |            |            | accgagggca               |      |
|            |            |            |            |            | cagaaggtgg               |      |
|            |            |            |            |            | aacatcaccg               |      |
|            |            |            |            |            | ttccgtccgg               |      |
|            |            |            |            |            | gtggtgaaga               |      |
|            |            |            |            |            | cgccagctgc               |      |
|            |            |            |            |            | cagcagcacc               |      |
|            |            |            |            |            | gccgtggagc               |      |
|            |            |            |            |            | tggatggagt               |      |
|            |            |            |            |            | gagagccaga               |      |
| accaycayga | yaayaacgag | cayyayccyc | cggagetgga | caagigggcc | agcctgtgga               | 7#0U |

|            |            |            |            |            |            | 2542 |
|------------|------------|------------|------------|------------|------------|------|
| actggttcaa | catcaccaac | tggctgtggt | gaggatccag | atctgctgtg | ccttctagtt | 3540 |
| gccagccatc | tgttgtttgc | ccctccccg  | tgccttcctt | gaccctggaa | ggtgccactc | 3600 |
| ccactgtcct | ttcctaataa | aatgaggaaa | ttgcatcgca | ttgtctgagt | aggtgtcatt | 3660 |
| ctattctggg | gggtggggtg | gggcaggaca | gcaaggggga | ggattgggaa | gacaatagca | 3720 |
| ggcatgctgg | ggatgcggtg | ggctctatgg | gtacccaggt | gctgaagaat | tgacccggtt | 3780 |
| cctcctgggc | cagaaagaag | caggcacatc | cccttctctg | tgacacaccc | tgtccacgcc | 3840 |
| cctggttctt | agttccagcc | ccactcatag | gacactcata | gctcaggagg | gctccgcctt | 3900 |
| caatcccacc | cgctaaagta | cttggagcgg | tctctccctc | cctcatcagc | ccaccaaacc | 3960 |
| aaacctagcc | tccaagagtg | ggaagaaatt | aaagcaagat | aggctattaa | gtgcagaggg | 4020 |
| agagaaaatg | cctccaacat | gtgaggaagt | aatgagagaa | atcatagaat | tttaaggcca | 4080 |
| tgatttaagg | ccatcatggc | cttaatcttc | cgcttcctcg | ctcactgact | cgctgcgctc | 4140 |
| ggtcgttcgg | ctgcggcgag | cggtatcagc | tcactcaaag | gcggtaatac | ggttatccac | 4200 |
| agaatcaggg | gataacgcag | gaaagaacat | gtgagcaaaa | ggccagcaaa | aggccaggaa | 4260 |
| ccgtaaaaag | gccgcgttgc | tggcgttttt | ccataggctc | cgccccctg  | acgagcatca | 4320 |
| caaaaatcga | cgctcaagtc | agaggtggcg | aaacccgaca | ggactataaa | gataccaggc | 4380 |
| gtttccccct | ggaagctccc | tegtgegete | tcctgttccg | accctgccgc | ttaccggata | 4440 |
| cctgtccgcc | tttctccctt | cgggaagcgt | ggcgctttct | catagctcac | gctgtaggta | 4500 |
| tctcagttcg | gtgtaggtcg | ttcgctccaa | gctgggctgt | gtgcacgaac | ccccgttca  | 4560 |
| gcccgaccgc | tgcgccttat | ccggtaacta | tcgtcttgag | tccaacccgg | taagacacga | 4620 |
| cttatcgcca | ctggcagcag | ccactggtaa | caggattagc | agagcgaggt | atgtaggcgg | 4680 |
| tgctacagag | ttcttgaagt | ggtggcctaa | ctacggctac | actagaagaa | cagtatttgg | 4740 |
| tatctgcgct | ctgctgaagc | cagttacctt | cggaaaaaga | gttggtagct | cttgatccgg | 4800 |
| caaacaaacc | accgctggta | gcggtggttt | ttttgtttgc | aagcagcaga | ttacgcgcag | 4860 |
| aaaaaaagga | tctcaagaag | atcctttgat | cttttctacg | gggtctgacg | ctcagtggaa | 4920 |
| cqaaaactca | cgttaaggga | ttttggtcat | gagattatca | aaaaggatct | tcacctagat | 4980 |
| ccttttaaat | taaaaatgaa | gttttaaatc | aatctaaagt | atatatgagt | aaacttggtc | 5040 |
| tgacagttac | caatgcttaa | tcaqtqaqqc | acctatctca | gcgatctgtc | tatttcgttc | 5100 |
| atccatagtt | gcctgactcg | qqqqqqqqq  | gcgctgaggt | ctgcctcgtg | aagaaggtgt | 5160 |
| tactaactca | taccaggect | gaatcqcccc | atcatccagc | cagaaagtga | gggagccacg | 5220 |
| gttgatgaga | gctttgttgt | aggtggacca | gttggtgatt | ttgaactttt | gctttgccac | 5280 |
| ggaacggtct | gcgttgtcgg | gaagatgcgt | gatctgatcc | ttcaactcag | caaaagttcg | 5340 |
| atttattcaa | caaagccgcc | gtcccgtcaa | gtcagcgtaa | tgctctgcca | gtgttacaac | 5400 |
| caattaacca | attctgatta | gaaaaactca | tcgagcatca | aatgaaactg | caatttattc | 5460 |
| atatcaggat | tatcaatacc | atatttttqa | aaaaqccgtt | tctgtaatga | aggagaaaac | 5520 |
| tcaccgagge | agttccatag | gatggcaaga | tcctggtatc | ggtctgcgat | tccgactcgt | 5580 |
| ccaacatcaa | tacaacctat | taatttcccc | tcgtcaaaaa | taaggttatc | aagtgagaaa | 5640 |
| tcaccatgag | tgacgactga | atccggtgag | aatggcaaaa | gcttatgcat | ttctttccag | 5700 |
| acttottcaa | caggccagcc | attacqctcq | tcatcaaaat | cactcgcatc | aaccaaaccg | 5760 |
| ttattcattc | gtgattgcgc | ctgagcgaga | cqaaatacgc | gategetgtt | aaaaggacaa | 5820 |
| ttacaaacaq | gaatcgaatg | caaccqqcqc | aggaacactg | ccagcgcatc | aacaatattt | 5880 |
| tcacctgaat | caggatattc | ttctaatacc | tggaatgctg | ttttcccggg | gatcgcagtg | 5940 |
|            | atgcatcatc |            |            |            |            |      |
|            | gccagtttag |            |            |            |            |      |
| ttqccatqtt | tcagaaacaa | ctctqqcqca | tegggettee | catacaatcg | atagattgtc | 6120 |
| gcacctgatt | gcccgacatt | atcqcqaqcc | catttatacc | catataaatc | agcatccatg | 6180 |
| ttggaattta | atcgcggcct | cgagcaagac | atttcccatt | gaatatggct | cataacaccc | 6240 |
| cttqtattac | tgtttatgta | agcagacagt | tttattqttc | atgatgatat | atttttatct | 6300 |
| tatacaatat | aacatcagag | attttgagac | acaacqtqqc | tttcccccc  | ccccattat  | 6360 |
| tgaagcattt | atcagggtta | ttgtctcatg | ageggataca | tatttqaatq | tatttagaaa | 6420 |
| aataaacaaa | taggggttcc | gcgcacattt | ccccgaaaaq | tgccacctga | cgtctaagaa | 6480 |
| accattatta | tcatgacatt | aacctataaa | aataggcgta | tcacqaqqcc | ctttcgtc   | 6538 |
|            |            |            |            |            |            |      |

<211> 6541
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 31
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggggg tcagcggtg ttggcgggtg tcgggggtg tcgggggtg tcgggggtg tcgggggtg tcgggggtg tcgggggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcqqqtq 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgq 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cgqtaaatgg 420 cocqcetqqc tqaccqcca acqacccccq cccattqacq tcaataatqa cqtatqttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacqqtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teceegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt tittggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgtcette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcaqcq 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920 aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040 tgtggaagga ggccaccacc accetgetet gegccagega egccaaggee tacgacaceg 2100 aggtgcacaa cgtgtgggcc acceacgcet gcgtgcccac cgaccccaac ccccaggagg 2160 tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacqacatg gtggagcaga 2220 tgcacqaqqa catcatcagc ctgtgggacc aqagcctqaa qccctqcqtq aaqctqaccc 2280 ccctqtqcqt gagcctgaag tgcaccgacg ctagcaccag ctgcaacacc agcgtgatca 2340 cccaggeetg ccccaaggtg agettegage ccateceeaa ccaetactge gccccegeeg 2400 gettegecat cetgaagtge aaggacaaga agtteaacgg caagggeece tgcaccaacg 2460 tgagcaccgt gcagtgcacc cacggcatcc gccccgtggt gagcacccag ctgctgctga 2520 acggtagect ggccgaggag gaggtggtga tecgcagege taaettegee gacaaegeea 2580 aggtgateat cgtgcagetg aacgagageg tggagateaa etgeaceege eecaacaaca 2640 acaccegcaa gagcatecac ateggeceeg geegegeett etacaccace ggegagatea 2700 teggegaeat cegecaggee cactgeaace tgageegege caagtggaac gacacectga 2760 acaagatcgt gatcaagctg cgcgagcagt tcggcaacaa gaccatcgtg ttcaagcaca 2820

| gcagcggcgg | 002000000  | atrotoacco | acaccttcaa | ctacaacaac | gagttettet | 2880 |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            |            |      |
| actgcaacag |            |            |            |            |            |      |
| rgggcaaggc | catgtaegee |            | geggeeagat | cegeegeage | agcaacatca | 3060 |
| ccggcctgct | getgaceege | gacggcggca | acagcaacaa | cyagagegag | accetecgee | 3120 |
|            |            |            |            |            | aaggtggtga |      |
|            |            |            |            |            | gcccgccagc |      |
| tgctgagcgg | catcgtgcag | cagcagaaca | acctgctgcg | cgccatcgag | gcccagcagc | 3240 |
| acctgctgca | gctgaccgtg | tggggcatca | agcagctgca | ggcccgcacc | ctggccgtgg | 3300 |
| agcgctacct | gaaggaccag | cagcagctcg | agcagatctg | gaaccacacc | acctggatgg | 3360 |
| agtgggaccg | cgagatcaac | aactacacca | gcctgatcca | cagcctgatc | gaggagagcc | 3420 |
| agaaccagca | ggagaagaac | gagcaggagc | tgctggagct | ggacaagtgg | gccagcctgt | 3480 |
| ggaactggtt | caacatcacc | aactggctgt | ggtgaggatc | cagatetget | gtgccttcta | 3540 |
|            |            |            |            |            | gaaggtgcca |      |
|            |            |            |            |            | agtaggtgtc |      |
| attctattct | ggggggtggg | gtggggcagg | acagcaaggg | ggaggattgg | gaagacaata | 3720 |
|            |            |            |            |            | aattgacccg |      |
|            |            |            |            |            | ccctgtccac |      |
|            |            |            |            |            | agggctccgc |      |
|            |            |            |            |            | agcccaccaa |      |
|            |            |            |            |            | taagtgcaga |      |
|            |            |            |            |            | aattttaagg |      |
|            |            |            |            |            | actcgctgcg |      |
|            |            |            |            |            | tacggttatc |      |
|            |            |            |            |            | aaaaggccag |      |
|            |            |            |            |            | ctgacgagca |      |
|            |            |            |            |            | aaagatacca |      |
|            |            |            |            |            | cgcttaccgg |      |
|            |            |            |            |            | cacgctgtag |      |
|            |            |            |            |            | aaccccccgt |      |
|            |            |            |            |            | cggtaagaca |      |
|            |            |            |            |            | ggtatgtagg |      |
|            |            |            |            |            | gaacagtatt |      |
|            |            |            |            |            | gctcttgatc |      |
|            |            |            |            |            | agattacgcg |      |
|            |            |            |            |            | acgctcagtg |      |
|            |            |            |            |            | tetteaceta |      |
|            |            |            |            |            | agtaaacttg |      |
|            |            |            |            |            | gtctatttcg |      |
|            |            |            |            |            | gtgaagaagg |      |
|            |            |            |            |            | tgagggagcc |      |
|            |            |            |            |            | tttgctttgc |      |
|            |            |            |            |            | cagcaaaagt |      |
|            |            |            |            |            | ccagtgttac |      |
|            |            |            |            |            | ctgcaattta |      |
|            |            |            |            |            | tgaaggagaa |      |
|            |            |            |            |            | gattecgact |      |
|            |            |            |            |            | atcaagtgag |      |
|            |            |            |            |            | catttctttc |      |
|            |            |            |            |            | atcaaccaaa |      |
|            |            |            |            |            | gttaaaagga |      |
|            |            |            |            |            | atcaacaata |      |
|            |            |            |            |            | ggggatcgca |      |
|            |            |            |            |            | cggaagaggc |      |
|            |            |            |            |            | ggcaacgcta |      |
| actuations | atttcagae  | caactctggc | gcatconoct | toccatacaa | tcgatagatt | 6120 |
| Coccegucat | Jecesagaaa |            | 2233500    |            |            |      |

```
gtcgcacctg attgcccgac attatcgcga gcccatttat acccatataa atcagcatcc 6180 atgttggaat ttaatcgcgg cctcgagcaa gacgtttccc gttgaatatg gctcataaca 6240 ccccttgtat tactgtttat gtaagcagac agttttattg ttcatgatga tatattttta 6300 tcttgtgcaa tgtaacatca gagattttga gacacaacgt ggctttcccc cccccccat 6360 tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 6420 aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa 6480 gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag gccctttcgt 6540 c
```

```
<210> 32
<211>, 6577
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttqqcaqtac atctacqtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaceteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
```

teccegtgce aagagtgacg taagtacege ctatagacte tataggcaca cccetttggc 1080 tettatgcat getatactgt tettggettg gggcetatac accecegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tggggttattg accattattg accaetecec 1200 tattggtgac gatactttce attactaate cataacatgg ctetttgcca caactatect 1260 tattggctat atgccaatac tetgtcette agagactgac acggactetg tattttaca 1320 ggatggggte ceattatta tetacaaatt cacatataca accaegegt ceceegtgcc 1380 cgcagtttt attacacata gegtgggate teccaegegaa tetegggtac gtgtteegg 1440 catgggetet teteceggtag eggeggaget teccaeaeeg agecetggte ceatgectec 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggcac 1560 agaaaatgage gtggagattg ggetegecae getgaeeg atggagaget taaggcage 1680

tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040 tgtggaagga ggccaccacc accetgctct gcgccagcga cgccaaggcc tacgacaccg 2100

|            |            |            |            |            | ccccaggagg |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | gtggagcaga |      |
| tgcacgagga | catcatcagc | ctgtgggacc | agagcctgaa | gccctgcgtg | aagctgaccc | 2280 |
| ccctgtgcgt | gagcctgaag | tgcaccgacg | ctagcaagaa | ctgcagcttc | aacatcagca | 2340 |
| ccagcatccg | cggcaaggtg | cagaaggagt | acgccttctt | ctacaagctg | gacatcatcc | 2400 |
| ccatcgacaa | cgacaccacc | agctacagcc | tgaccagctg | caacaccagc | gtgatcaccc | 2460 |
| aggcctgccc | caaggtgagc | ttcgagccca | tccccaacca | ctactgcgcc | cccgccggct | 2520 |
|            |            |            |            |            | accaacgtga |      |
| gcaccgtgca | gtgcacccac | ggcatccgcc | ccgtggtgag | cacccagctg | ctgctgaacg | 2640 |
| gtagcctggc | cgaggaggag | gtggtgatcc | gcagcgctaa | cttcgccgac | aacgccaagg | 2700 |
| tgatcatcgt | gcagctgaac | gagagcgtgg | agatcaactg | cacccgcgct | agcgcccact | 2760 |
|            |            |            |            |            | aagctgcgcg |      |
| agcagttcgg | caacaagacc | atcgtgttca | agcacagcag | cggcggcgac | cccgagatcg | 2880 |
| tgacccacag | cttcaactgc | ggcggcgagt | tcttctactg | caacagcgct | agcctgccct | 2940 |
| gccgcatcaa | gcagatcatc | aacatgtggc | agaaggtggg | caaggccatg | tacgcccccc | 3000 |
| ccatcagcgg | ccagatccgc | tgcagcagca | acatcaccgg | cctgctgctg | acccgcgacg | 3060 |
| gcggcaacag | caacaacgag | agcgagatct | tccgtccggg | cggcggcgac | atgcgcgaca | 3120 |
| actggcgcag | cgagctgtac | aagtacaagg | tggtgaagat | cgagcccctg | ggcgtggccc | 3180 |
| ccaccaaggc | caagcttacc | gtccaggccc | gccagctgct | gagcggcatc | gtgcagcagc | 3240 |
| agaacaacct | gctgcgcgcc | atcgaggccc | agcagcacct | gctgcagctg | accgtgtggg | 3300 |
| gcatcaagca | gctgcaggcc | cgcaccctgg | ccgtggagcg | ctacctgaag | gaccagcagc | 3360 |
| agctcgagca | gatctggaac | cacaccacct | ggatggagtg | ggaccgcgag | atcaacaact | 3420 |
| acaccagect | gatccacagc | ctgatcgagg | agagccagaa | ccagcaggag | aagaacgagc | 3480 |
| aggagctgct | ggagctggac | aagtgggcca | gcctgtggaa | ctggttcaac | atcaccaact | 3540 |
| ggctgtggtg | aggatccaga | tctgctgtgc | cttctagttg | ccagccatct | gttgtttgcc | 3600 |
| cctccccgt  | gccttccttg | accctggaag | gtgccactcc | cactgtcctt | tcctaataaa | 3660 |
| atgaggaaat | tgcatcgcat | tgtctgagta | ggtgtcattc | tattctgggg | ggtggggtgg | 3720 |
| ggcaggacag | caagggggag | gattgggaag | acaatagcag | gcatgctggg | gatgcggtgg | 3780 |
| gctctatggg | tacccaggtg | ctgaagaatt | gacccggttc | ctcctgggcc | agaaagaagc | 3840 |
| aggcacatcc | ccttctctgt | gacacaccct | gtccacgccc | ctggttctta | gttccagccc | 3900 |
| cactcatagg | acactcatag | ctcaggaggg | ctccgccttc | aatcccaccc | gctaaagtac | 3960 |
| ttggagcggt | ctctccctcc | ctcatcagcc | caccaaacca | aacctagcct | ccaagagtgg | 4020 |
| gaagaaatta | aagcaagata | ggctattaag | tgcagaggga | gagaaaatgc | ctccaacatg | 4080 |
| tgaggaagta | atgagagaaa | tcatagaatt | ttaaggccat | gatttaaggc | catcatggcc | 4140 |
| ttaatcttcc | gcttcctcgc | tcactgactc | gctgcgctcg | gtcgttcggc | tgcggcgagc | 4200 |
| ggtatcagct | cactcaaagg | cggtaatacg | gttatccaca | gaatcagggg | ataacgcagg | 4260 |
| aaagaacatg | tgagcaaaag | gccagcaaaa | ggccaggaac | cgtaaaaagg | ccgcgttgct | 4320 |
| ggcgttttc  | cataggctcc | gcccccctga | cgagcatcac | aaaaatcgac | gctcaagtca | 4380 |
| gaggtggcga | aacccgacag | gactataaag | ataccaggcg | tttccccctg | gaagctccct | 4440 |
| cgtgcgctct | cctgttccga | ccctgccgct | taccggatac | ctgtccgcct | ttctcccttc | 4500 |
| gggaagcgtg | gcgctttctc | atagctcacg | ctgtaggtat | ctcagttcgg | tgtaggtcgt | 4560 |
| tegetecaag | ctgggctgtg | tgcacgaacc | ccccgttcag | cccgaccgct | gcgccttatc | 4620 |
| cggtaactat | cgtcttgagt | ccaacccggt | aagacacgac | ttatcgccac | tggcagcagc | 4680 |
| cactggtaac | aggattagca | gagcgaggta | tgtaggcggt | gctacagagt | tcttgaagtg | 4740 |
| gtggcctaac | tacggctaca | ctagaagaac | agtatttggt | atctgcgctc | tgctgaagcc | 4800 |
| agttaccttc | ggaaaaagag | ttggtagctc | ttgatccggc | aaacaaacca | ccgctggtag | 4860 |
| cggtggtttt | tttgtttgca | agcagcagat | tacgcgcaga | aaaaaaggat | ctcaagaaga | 4920 |
| tcctttgatc | ttttctacgg | ggtctgacgc | tcagtggaac | gaaaactcac | gttaagggat | 4980 |
|            |            |            |            |            | aaaaatgaag |      |
|            |            |            |            |            | aatgcttaat |      |
|            |            |            |            |            | cctgactcgg |      |
|            |            |            |            |            | accaggcctg |      |
|            |            |            |            |            | ctttgttgta |      |
|            |            |            |            |            | cgttgtcggg |      |
| aagatgcgtg | atctgatcct | tcaactcagc | aaaagttcga | tttattcaac | aaagccgccg | 5400 |

```
tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa ttctgattag 5460
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5520
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5580
atggcaagat cetggtateg gtetgegatt cegactegte caacateaat acaacetatt 5640
aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5700
teeggtgaga atggcaaaag ettatgeatt tettteeaga ettgtteaac aggceageea 5760
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 5820
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 5880
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 5940
tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tgcatcatca 6000
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6060
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6120
tetggegeat egggetteee atacaatega tagattgteg cacetgattg ecegacatta 6180
tegegagece atttatacce atataaatca geatecatgt tggaatttaa tegeggeete 6240
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6300
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6360
ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta tcagggttat 6420
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat agggqttccq 6480
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6540
acctataaaa ataggcgtat cacgaggccc tttcgtc
                                                                  6577
```

<210> 33
<211> 6538
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HTV genes

<400> 33 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagacggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teageggtt 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cqqtaaatqq 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcategc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcqtaacaa 840 ctccgccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgacg taagtacege ctatagacte tataggeaca eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acccccgett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactitcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagaetgae aeggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacqccqt cccccqtqcc 1380

|            |            |            |            |              | gtgttccgga |      |
|------------|------------|------------|------------|--------------|------------|------|
|            |            |            |            |              | ccatgcctcc |      |
|            |            |            |            |              | acttaggcac |      |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg   | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact   | taaggcagcg | 1680 |
|            |            |            |            |              | ggtaactccc |      |
|            |            |            |            |              | tgctgccgcg |      |
|            |            |            |            |              | gggtcttttc |      |
|            |            |            |            |              | atgcgcgtga |      |
|            |            |            |            |              | ctgctgggca |      |
|            |            |            |            |              | ggcgtgcccg |      |
|            |            |            |            |              | tacgacaccg |      |
|            |            |            |            |              | ccccaggagg |      |
|            |            |            |            |              |            |      |
|            |            |            |            |              | gtggagcaga |      |
|            |            |            |            |              | aagctgaccc |      |
|            |            |            |            |              | aacagcagca |      |
|            |            |            |            |              | agcaccagct |      |
|            |            |            |            |              | atccccaacc |      |
|            |            |            |            |              | ttcaacggca |      |
|            |            |            |            |              | cccgtggtga |      |
| gcacccagct | gctgctgaac | ggtagcctgg | ccgaggagga | ggtggtgatc   | cgcagcgcta | 2640 |
| acttcgccga | caacgccaag | gtgatcatcg | tgcagctgaa | cgagagcgtg   | gagatcaact | 2700 |
| gcacccgcgc | tagcgcccac | tgcaacctga | gccgcgccaa | gtggaacgac   | accctgaaca | 2760 |
| agatcgtgat | caagetgege | gagcagttcg | gcaacaagac | catcgtgttc   | aagcacagca | 2820 |
|            |            |            |            |              | ttcttctact |      |
|            |            |            |            |              | cagaaggtgg |      |
|            |            |            |            |              | aacatcaccg |      |
|            |            |            |            |              | ttccgtccgg |      |
|            |            |            |            |              | gtggtgaaga |      |
|            |            |            |            |              | cgccagctgc |      |
|            |            |            |            |              | cagcagcacc |      |
|            |            |            |            |              | gccgtggagc |      |
|            |            |            |            |              | tggatggagt |      |
|            |            |            |            |              | gagagccaga |      |
|            |            |            |            |              |            |      |
|            |            |            |            |              | agcctgtgga |      |
|            |            |            |            |              | ccttctagtt |      |
|            |            |            |            |              | ggtgccactc |      |
|            |            |            |            |              | aggtgtcatt |      |
|            |            |            |            |              | gacaatagca |      |
|            |            |            |            |              | tgacccggtt |      |
|            |            |            |            |              | tgtccacgcc |      |
|            |            |            |            |              | gctccgcctt |      |
|            |            |            |            |              | ccaccaaacc |      |
|            |            |            |            |              | gtgcagaggg |      |
| agagaaaatg | cctccaacat | gtgaggaagt | aatgagagaa | atcatagaat   | tttaaggcca | 4080 |
|            |            |            |            |              | cgctgcgctc |      |
| ggtcgttcgg | ctgcggcgag | cggtatcagc | tcactcaaag | gcggtaatac   | ggttatccac | 4200 |
| agaatcaggg | gataacgcag | gaaagaacat | gtgagcaaaa | ggccagcaaa   | aggccaggaa | 4260 |
|            |            |            |            |              | acgagcatca |      |
|            |            |            |            |              | gataccaggc |      |
|            |            |            |            |              | ttaccggata |      |
|            |            |            |            |              | gctgtaggta |      |
|            |            |            |            |              | ccccgttca  |      |
|            |            |            |            |              | taagacacga |      |
|            |            |            |            |              | atgtaggcgg |      |
| 2002203004 |            |            |            | יונבייפייכיי | -555-35    |      |

```
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4740
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4800
caaacaaacc accoctoota ocooptoottt ttttotttoc aagcagcaga ttacgcgcag 4860
aaaaaaagga totcaagaag atcotttgat ottttctacg gggtctgacg otcagtggaa 4920
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 4980
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5040
tqacaqttac caatgcttaa tcaqtqaqqc acctatctca gcgatctgtc tatttcgttc 5100
atccatagtt gcctgactcg ggggggggg gcgctgaggt ctgcctcgtg aagaaggtgt 5160
tgctgactca taccaggect gaategeece atcatecage cagaaagtga gggagecaeg 5220
gttqatqaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5280
qqaacqqtct qcqttqtcqq qaaqatqcqt gatctqatcc ttcaactcag caaaagttcg 5340
atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5400
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttattc 5460
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5520
tcaccgagge agttccatag gatggcaaga tcctggtatc ggtctgcgat tccgactcgt 5580
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5640
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5700
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 5760
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 5820
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 5880
tracetgaat caggatatte ttetaatace tggaatgetg tttteceggg gategeagtg 5940
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6000
aatteegtea geeagtttag tetgaceate teatetgtaa cateattgge aacgetacet 6060
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6120
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6180
ttqqaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6240
cttqtattac tqtttatqta agcagacagt tttattqttc atgatgatat atttttatct 6300
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc cccccattat 6360
tqaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6420
aataaacaaa taqqqqttcc qcqcacattt ccccqaaaag tqccacctga cqtctaagaa 6480
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc
```

```
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 34
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
```

<210> 34 <211> 6460

| cataaataaa | cgtggatagc | aatttaaata | aggggattt  | casatctcc  | accccattca | 780  |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            |            |      |
|            | agtttgttt  |            |            |            |            |      |
|            | ttgacgcaaa |            |            |            |            |      |
|            | gtgaaccgtc |            |            |            |            |      |
|            | cgggaccgat |            |            |            |            |      |
|            | aagagtgacg |            |            |            |            |      |
|            | gctatactgt |            |            |            |            |      |
|            | tatagcttag |            |            |            |            |      |
|            | gatactttcc |            |            |            |            |      |
|            | atgccaatac |            |            |            |            |      |
|            | ccatttatta |            |            |            |            |      |
|            | attaaacata |            |            |            |            |      |
| catgggctct | tctccggtag | cggcggagct | tccacatccg | agccctggtc | ccatgcctcc | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact | taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
|            | tgttaacggt |            |            |            |            |      |
|            | gacataatag |            |            |            |            |      |
|            | gtcgtcgaca |            |            |            |            |      |
|            | ccagcacctg |            |            |            |            |      |
|            | ctgcagcgcc |            |            |            |            |      |
|            | ggccaccacc |            |            |            |            |      |
|            | cgtgtgggcc |            |            |            |            |      |
|            | gaacgtgacc |            |            |            |            |      |
|            | catcatcagc |            |            |            |            |      |
|            | gagcctgaag |            |            |            |            |      |
|            | ccccaaggtg |            |            |            |            |      |
|            | cctgaagtgc |            |            |            |            |      |
|            | gcagtgcacc |            |            |            |            |      |
|            | ggccgaggag |            |            |            |            |      |
|            | cgtgcagctg |            |            |            |            |      |
|            | gagccgcgcc |            |            |            |            |      |
|            | cggcaacaag |            |            |            |            |      |
|            | cagcttcaac |            |            |            |            |      |
|            | caagcagatc |            |            |            |            |      |
|            | cggccagatc |            |            |            |            |      |
|            | cagcaacaac |            |            |            |            |      |
|            | cagcgagctg |            |            |            |            |      |
|            | ggccaagctt |            |            |            |            |      |
|            | cctgctgcgc |            |            |            |            |      |
|            | gcagctgcag |            |            |            |            |      |
|            | gcagatctgg |            |            |            |            |      |
|            | cctgatccac |            |            |            |            |      |
|            | gctggagctg |            |            |            |            |      |
|            | gtgaggatcc |            |            |            |            |      |
|            | cgtgccttcc |            |            |            |            |      |
|            | aattgcatcg |            |            |            |            |      |
|            | cagcaagggg |            |            |            |            |      |
|            | gggtacccag |            |            |            |            |      |
|            | tecettete  |            |            |            |            |      |
|            | aggacactca |            |            |            |            |      |
|            | ggtctctccc |            |            |            |            |      |
|            | ttaaagcaag |            |            |            |            |      |
|            | gtaatgagag |            |            |            |            |      |
|            |            | -          |            | -          | _          |      |

```
geottaatet teegetteet egeteactga etegetgege teggtegtte ggetgeggeg 4080
ageggtatea geteacteaa aggeggtaat aeggttatee acagaateag gggataaege 4140
aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt 4200
gctggcgttt ttccataggc tccgccccc tgacgagcat cacaaaaatc gacgctcaag 4260
tcaqaqqtqq cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc 4320
cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc 4380
ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt 4440
cgttcgctcc aagctgggct gtgtgcacga accccccgtt cagcccgacc gctgcgcctt 4500
atcoggtaac tatogtottg agtocaacco ggtaagacac gacttatogc cactggcage 4560
agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa 4620
gtggtggcct aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa 4680
qccaqttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg 4740
tagcggtggt ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga 4800
agateetttg atettteta eggggtetga egeteagtgg aacgaaaact caegttaagg 4860
gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg 4920
aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt 4980
aatcaqtqaq qcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact 5040
cgggggggg gggcgctgag gtctgcctcg tgaagaaggt gttgctgact cataccaggc 5100
ctgaatcgcc ccatcatcca gccagaaagt gagggagcca cggttgatga gagctttgtt 5160
qtaqqtqqac cagttggtga ttttgaactt ttgctttgcc acggaacggt ctgcgttgtc 5220
gggaagatgc gtgatctgat ccttcaactc agcaaaagtt cgatttattc aacaaagccg 5280
ccqtcccqtc aagtcagcgt aatgctctgc cagtgttaca accaattaac caattctgat 5340
tagaaaaact catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata 5400
ccatattttt gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat 5460
aggatggcaa gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct 5520
attaatttcc cctcqtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact 5580
quatccggtg agaatggcaa aagcttatgc atttctttcc agacttgttc aacaggccag 5640
ccattacgct cgtcatcaaa atcactcgca tcaaccaaac cgttattcat tcgtgattgc 5700
gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac aattacaaac aggaatcgaa 5760
tgcaaccggc gcaggaacac tgccagcgca tcaacaatat tttcacctga atcaggatat 5820
tottotaata cotggaatgo tgttttcccg gggatcgcag tggtgagtaa ccatgcatca 5880
tcaggagtac ggataaaatg cttgatggtc ggaagaggca taaattccgt cagccagttt 5940
agtotgacca totoatotgt aacatoattg gcaacgotac otttgccatg tttcagaaac 6000
aactotggog catogggott occatacaat ogatagattg togcacotga ttgccogaca 6060
ttatcgcgag cccatttata cccatataaa tcagcatcca tgttggaatt taatcgcggc 6120
ctcgagcaag acgtttcccg ttgaatatgg ctcataacac cccttgtatt actgtttatg 6180
taagcagaca gttttattgt tcatgatgat atatttttat cttgtgcaat gtaacatcag 6240
agattttgag acacaacgtg gctttccccc ccccccatt attgaagcat ttatcagggt 6300
tattqtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt 6360
ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca 6420
ttaacctata aaaataggcg tatcacgagg ccctttcgtc
```

```
<210> 35
<211> 7987
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pAdApt containing HIV genes
<400> 35
ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttattt 60
```

qqattqaaqc caatatgata atgagggggt ggagtttgtg acgtggcgcg gggcgtggga 120

| WO 02/32943 | •          |            |            |            | PCT/US     | 01/25721 |
|-------------|------------|------------|------------|------------|------------|----------|
| acqqqqcqqq  | tgacgtagta | qtqtqqcqqa | agtgtgatgt | tgcaagtgtg | qcqqaacaca | 180      |
|             | ggatgtggca |            |            |            |            |          |
|             | gcgcggtttt |            |            |            |            |          |
|             | ttcgcgggaa |            |            |            |            |          |
|             | cgtaatattt |            |            |            |            |          |
|             | tttttctcag |            |            |            |            |          |
|             | acgtaccagt |            |            |            |            |          |
|             | tgtacattta |            |            |            |            |          |
|             | gttattaata |            |            |            |            |          |
|             | ttacataact |            |            |            |            |          |
|             | cgtcaataat |            |            |            |            |          |
|             | gggtggagta |            |            |            |            |          |
|             | gtacgccccc |            |            |            |            |          |
|             | tgaccttatg |            |            |            |            |          |
|             | tggtgatgcg |            |            |            |            |          |
|             | ttccaagtct |            |            |            |            |          |
|             | actttccaaa |            |            |            |            |          |
|             | ggtgggaggt |            |            |            |            |          |
|             | atccacgctg |            |            |            |            |          |
|             | gtcgacacgt |            |            |            |            |          |
|             | gcacctgtgg |            |            |            |            |          |
|             | cagegeeace |            |            |            |            |          |
|             | caccaccacc |            |            |            |            |          |
|             | gtgggccacc |            |            |            |            |          |
|             | cgtgaccgag |            |            |            |            |          |
|             | catcagcctg |            |            |            |            |          |
|             | cctgaagtgc |            |            |            |            |          |
|             | caaggtgcag |            |            |            |            |          |
|             | caccaccagc |            |            |            |            |          |
|             | ggtgagcttc |            |            |            |            |          |
|             | gtgcaaggac |            |            |            |            |          |
|             | cacccacggc |            |            |            |            |          |
|             | ggaggaggtg |            |            |            |            |          |
|             | gctgaacgag |            |            |            |            |          |
| gcaagagcat  | ccacatcggc | cccggccgcg | ccttctacac | caccggcgag | atcatcggcg | 2220     |
| acatccgcca  | ggcccactgc | aacctgagcc | gcgccaagtg | gaacgacacc | ctgaacaaga | 2280     |
|             | gctgcgcgag |            |            |            |            |          |
| geggegaece  | cgagatcgtg | acccacagct | tcaactgcgg | cggcgagttc | ttctactgca | 2400     |
| acagcaccca  | gctgttcaac | agcacctggt | tcaacagcac | ctggagcacc | gagggcagca | 2460     |
| acaacaccga  | gggcagcgac | accatcaccc | tgccctgccg | catcaagcag | atcatcaaca | 2520     |
|             | ggtgggcaag |            |            |            |            |          |
| gcagcaacat  | caccggcctg | ctgctgaccc | gcgacggcgg | caacagcaac | aacgagagcg | 2640     |
|             | tccgggcggc |            |            |            |            |          |
|             | gaagatcgag |            |            |            |            |          |
|             | gctgctgagc |            |            |            |            |          |
|             | gcacctgctg |            |            |            |            |          |
|             | ggagcgctac |            |            |            |            |          |
|             | ggagtgggac |            |            |            |            |          |
|             | ccagaaccag |            |            |            |            |          |
|             | gtggaactgg |            |            |            |            |          |
|             | tagttgccag |            |            |            |            |          |
|             | cactcccact |            |            |            |            |          |
|             | tcattctatt |            |            |            |            |          |
|             | tagcaggcat |            |            |            |            |          |
| taacttcgta  | taatgtatgc | tatacgaagt | tataagatct | gtactgaaat | gtgtgggcgt | 3420     |

| gacttaagga  | tgggaaagaa | tatataaggt | gggggtctta | tataatttta | tatctgtttt | 3480 |
|-------------|------------|------------|------------|------------|------------|------|
|             |            |            |            |            | gagctcatat |      |
| ttgacaacgc  | gcatgcccc  | atagaccaga | atacatcaaa | atataataaa | ctccagcatt | 3600 |
|             |            |            |            |            | cgtgtctgga |      |
| acaccattaa  | agactgcagc | chacacaca  | acttcaacca | ctocaoccac | cgcccgcggg | 3720 |
| attotoacto  | actitectit | cctgagcccg | cttocaagea | atacaacttc | ccgttcatcc | 3780 |
| accordata   | acaagttgac | aactettta  | gcacaattgg | attetttgac | ccgggaactt | 3840 |
| aatataatt   | ctcagcagac | attagatcta | caccaacaa  | tttctaccct | gaaggcttcc | 3900 |
| tacactaca   | atacaattta | aaacataaat | aaaaaaccad | actctgtttg | gatttggatc | 3960 |
|             |            |            |            |            | ccgggaccag |      |
|             |            |            |            |            | gtgactctgg |      |
|             |            |            |            |            | ctgcagagct |      |
|             |            |            |            |            | ggcgtggtgc |      |
|             |            |            |            |            | gtaagtgttt |      |
|             |            |            |            |            | cttggactgt |      |
|             |            |            |            |            | gtgcagaacc |      |
| accectagge  | tatatacaat | acacttagaa | aatttataat | gacccatgee | aggaaatgcg | 4440 |
| terangement | tagaaaaaaa | cttctcacct | ccaacattt  | ccatacatta | gtccataatg | 4500 |
| cygaagaacc  | accesace   | gagagatta  | accasastat | ttctgggatc | actaacgtca | 4560 |
|             |            |            |            |            |            |      |
|             |            |            |            |            | gcggagggtg |      |
|             |            |            |            |            | acagatttgc |      |
|             |            |            |            |            | gatgaagaaa |      |
|             |            |            |            |            | cagctgcgac |      |
| ttaccgcagc  | eggegggeee | gtaaattata | cctattaccg | getgeaacty | gtagttaaga | 4000 |
|             |            |            |            |            | gtccctgact |      |
|             |            |            |            |            | tagcagttct |      |
| tgcaaggaag  | caaagttttt | caacggtttg | agaeegeeeg | cegtaggeat | gcttttgagc | 5040 |
|             |            |            |            |            | ggcatctcga |      |
|             |            |            |            |            | agtagtcggt |      |
|             |            |            |            |            | gtcagcgtag |      |
|             |            |            |            |            | cgcttgaggc |      |
|             |            |            |            |            | aggtagcatt |      |
|             |            |            |            |            | agettgeeet |      |
|             |            |            |            |            | ttgggcgcga |      |
|             |            |            |            |            | gtctcgcatt |      |
|             |            |            |            |            | ccatgctttt |      |
|             |            |            |            |            | acgaaaaggc |      |
|             |            |            |            |            | ccgcggtcct |      |
|             |            |            |            |            | agcacgaagg |      |
|             |            |            |            |            | tccagggtgt |      |
|             |            |            |            |            | gtgtaggcca |      |
|             |            |            |            |            | tegteeteae |      |
|             |            |            |            |            | cgaggctgga |      |
|             |            |            |            |            | gcgttgcagg |      |
|             |            |            |            |            | cagcaaaagg |      |
|             |            |            |            |            | cccctgacg  |      |
|             |            |            |            |            | ctataaagat |      |
|             |            |            |            |            | ctgccgctta |      |
|             |            |            |            |            | agctcacgct |      |
|             |            |            |            |            | cacgaacccc |      |
|             |            |            |            |            | aacccggtaa |      |
|             |            |            |            |            | gcgaggtatg |      |
|             |            |            |            |            | agaaggacag |      |
| tatttggtat  | ctgcgctctg | ctgaagccag | ttaccttcgg | aaaaagagtt | ggtagctctt | 6660 |
| gatccggcaa  | acaaaccacc | gerggrageg | gtggttttt  | cgtttgcaag | cagcagatta | 6/20 |

```
cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 6780
aqtqqaacqa aaactcacqt taaqqqattt tqqtcatqaq attatcaaaa aggatcttca 6840
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 6900
cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 6960
ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct 7020
taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt 7080
tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat 7140
cogcetecat coaqtetatt aattqttqcc qqqaaqctaq aqtaaqtaqt tegecagtta 7200
ataqtttqcq caacqttqtt gccattgctg caggcatcgt ggtgtcacgc tcgtcgtttg 7260
gtatggette atteagetee ggtteecaae gateaaggeg agttacatga tececcatgt 7320
tqtqcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg 7380
cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg 7440
taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc 7500
qqcqaccqaq ttgctcttgc ccggcgtcaa cacgggataa taccgcgcca catagcagaa 7560
ctttaaaaqt qctcatcatt qqaaaacqtt cttcqqqqcg aaaactctca aggatcttac 7620
cqctqttqaq atccaqttcg atgtaaccca ctcgtgcacc caactgatct tcagcatctt 7680
ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg 7740
gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa 7800
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 7860
aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca 7920
ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtcttcaag 7980
aattgtt
```

<210> 36
<211> 7948
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pAdApt containing HIV genes
<400> 36
thanktages generates to the tagency of tagency of the tagency of tagency o

ttaattaacc gcaattotoa tgtttgacag ottatoatoa toaataatat acottatttt 60 ggattgaagc caatatgata atgaggggt ggagtttgtg acgtggcgcg gggcgtggga 120 acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180 tgtaagogac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240 gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300 tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360 acteatageg egtaatattt gtetagggee geggggaett tgacegttta egtggagaet 420 cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480 tatagtcagt acgtaccagt gcactggcct agageggccc cattgcatac gttgtatcca 540 tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600 ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660 qaqttccqcq ttacataact tacqqtaaat gqcccqcctq qctqaccqcc caacqacccc 720 cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780 tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840 catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900 gcccagtaca tgaccttatg ggactttect acttggcagt acatetacgt attagtcate 960 gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020 tcacggggat ttccaaqtct ccaccccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080 aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140 aggogtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200 tggagacgcc atccacqctg ttttgacctc catagaagac accgggaccg atccaqcctc 1260

| cgtcaccgtc | gtcgacacgt | gtgatcagat | atcgcggccg | ctctagaatg               | cgcgtgaagg | 1320 |
|------------|------------|------------|------------|--------------------------|------------|------|
|            |            |            |            | caccatgctg               |            |      |
|            |            |            |            | gtactacggc               |            |      |
|            |            |            |            | caaggcctac               |            |      |
|            |            |            |            | ccccaacccc               |            |      |
|            |            |            |            | cgacatggtg               |            |      |
|            |            |            |            | ctgcgtgaag               |            |      |
|            |            |            |            | caacaccaac               |            |      |
|            |            |            |            | cttcgctagc               |            |      |
|            |            |            |            | cgagcccatc               |            |      |
|            |            |            |            | caagaagttc               |            |      |
|            |            |            |            | catccgcccc               |            |      |
|            |            |            |            | ggtgatccgc               |            |      |
|            |            |            |            | gagcgtggag               |            |      |
|            |            |            |            | ccccggccgc               |            |      |
|            |            |            |            | caacctgage               |            |      |
|            |            |            |            | gcagttcggc               |            |      |
|            |            |            |            | gacccacage               |            |      |
|            |            |            |            | cagcacctgg               |            |      |
|            |            |            |            | caccatcacc               |            |      |
|            |            |            |            | ggccatgtac               |            |      |
|            |            |            |            | gctgctgacc               |            |      |
|            |            |            |            | cggcgacatg               |            |      |
|            |            |            |            | gcccctgggc               |            |      |
|            |            |            |            | cggcatcgtg               |            |      |
|            |            |            |            |                          |            |      |
|            |            |            |            | gcagctgacc               |            |      |
|            |            |            |            | cctgaaggac               |            |      |
|            |            |            |            | ccgcgagatc               |            |      |
|            |            |            |            | gcaggagaag<br>gttcaacatc |            |      |
|            |            |            |            |                          |            |      |
|            |            |            |            | gccatctgtt               |            |      |
|            |            |            |            | tgtcctttcc<br>tctggggggt |            |      |
|            |            |            |            | tgctggggat               |            |      |
|            |            |            |            |                          |            |      |
|            |            |            |            | ctatacgaag               |            |      |
|            |            |            |            | atatataagg               |            |      |
|            |            |            |            | tgagcaccaa               |            |      |
|            |            |            |            | catgggccgg               |            |      |
|            |            |            |            | ccgcaaactc               |            |      |
|            |            |            |            | cctccgccgc               |            |      |
|            |            |            |            | tectgagece               |            |      |
|            |            |            |            | cggctctttt               |            |      |
|            |            |            |            | tgttggatct               |            |      |
|            |            |            |            | aaaacataaa               |            |      |
|            |            |            |            | tttatttagg               |            |      |
|            |            |            |            | tcctgtgtat               |            |      |
|            |            |            |            | taagcccgtc               |            |      |
|            |            |            |            | tgtagatgat               |            |      |
|            |            |            |            | gcaagctgat               |            |      |
|            |            |            |            | atgggtgcat               |            |      |
|            |            |            |            | tcccagccat               |            |      |
|            |            |            |            | tgcacttggg               |            |      |
|            |            |            |            | ccttgtgacc               |            |      |
|            |            |            |            | cggcggcctg               |            |      |
| tttctgggat | cactaacgtc | atagttgtgt | tccaggatga | gatcgtcata               | ggccattttt | 4560 |

|            | aaaaaaaa   | aaaaaaataa | aatstsstaa | ttaastaaaa  | ~~~~~~~~   | 4620 |
|------------|------------|------------|------------|-------------|------------|------|
| tacktaccct | ggcggagggc | gecagaetge | ggtataatgg | cocaccegg   | cccaggggcg | 4620 |
| nagttaccct | cacagacceg | aaccatttaa | gerregager | cagacggggg  | gatcatgtct | 4740 |
|            |            |            |            |             | ggaagaaagc |      |
|            |            |            |            |             | acctattacc |      |
|            |            |            |            |             | gggggccact |      |
|            |            |            |            |             | aaggcgctcg |      |
|            |            |            |            |             | gagaccgtcc |      |
|            |            |            |            |             | cagctcggtc |      |
|            |            |            |            |             | gggcggcttt |      |
|            |            |            |            |             | ttccacgggc |      |
|            |            |            |            |             | ggctgcgcgc |      |
|            |            |            |            |             | tettegeeet |      |
|            |            |            |            |             | gcggcgtggc |      |
|            |            |            |            |             | agacttttga |      |
|            |            |            |            |             | gcgccgcagg |      |
|            |            |            |            |             | gggtcaaaaa |      |
|            |            |            |            |             | agccggtgtc |      |
|            |            |            |            |             | ggcctgtcct |      |
|            |            |            |            |             | acaaaggctc |      |
| gcgtccaggc | cagcacgaag | gaggctaagt | gggaggggta | gcggtcgttg  | tccactaggg | 5760 |
|            |            |            |            |             | aggaaggtga |      |
|            |            |            |            |             | taaaaggggg |      |
|            |            |            |            |             | agctgttggg |      |
|            |            |            |            |             | tccggcggca |      |
|            |            |            |            |             | catcagggac |      |
| agcttcaagg | ccagcaaaag | gccaggaacc | gtaaaaaggc | cgcgttgctg  | gcgtttttcc | 6120 |
|            |            |            |            |             | aggtggcgaa |      |
|            |            |            |            |             | gtgcgctctc |      |
|            |            |            |            |             | ggaagcgtgg |      |
| cgctttctca | tagctcacgc | tgtaggtatc | tcagttcggt | gtaggtcgtt  | cgctccaagc | 6360 |
| tgggctgtgt | gcacgaaccc | cccgttcagc | ccgaccgctg | cgccttatcc  | ggtaactatc | 6420 |
|            |            |            |            |             | actggtaaca |      |
|            |            |            |            |             | tggcctaact |      |
| acggctacac | tagaaggaca | gtatttggta | tetgegetet | gctgaagcca  | gttaccttcg | 6600 |
| gaaaaagagt | tggtagctct | tgatccggca | aacaaaccac | cgctggtagc  | ggtggttttt | 6660 |
| ttgtttgcaa | gcagcagatt | acgcgcagaa | aaaaaggatc | tcaagaagat  | cctttgatct | 6720 |
| tttctacggg | gtctgacgct | cagtggaacg | aaaactcacg | ttaagggatt  | ttggtcatga | 6780 |
| gattatcaaa | aaggatcttc | acctagatcc | ttttaaatta | aaaatgaagt  | tttaaatcaa | 6840 |
| tctaaagtat | atatgagtaa | acttggtctg | acagttacca | atgcttaatc  | agtgaggcac | 6900 |
| ctatctcagc | gatctgtcta | tttcgttcat | ccatagttgc | ctgactcccc  | gtcgtgtaga | 6960 |
| taactacgat | acgggagggc | ttaccatctg | gccccagtgc | tgcaatgata  | ccgcgagacc | 7020 |
|            |            |            |            |             | gccgagcgca |      |
| gaagtggtcc | tgcaacttta | tccgcctcca | tccagtctat | taattgttgc  | cgggaagcta | 7140 |
|            |            |            |            |             | gcaggcatcg |      |
|            |            |            |            |             | cgatcaaggc |      |
| gagttacatg | atcccccatg | ttgtgcaaaa | aagcggttag | ctccttcggt. | cctccgatcg | 7320 |
| ttgtcagaag | taagttggcc | gcagtgttat | cactcatggt | tatggcagca  | ctgcataatt | 7380 |
| ctcttactgt | catgccatcc | gtaagatgct | tttctgtgac | tggtgagtac  | tcaaccaagt | 7440 |
|            |            |            |            |             | acacgggata |      |
|            |            |            |            |             | tcttcggggc |      |
|            |            |            |            |             | actcgtgcac |      |
|            |            |            |            |             | aaaacaggaa |      |
|            |            |            |            |             | ctcatactct |      |
|            |            |            |            |             | ggatacatat |      |
|            |            |            |            |             | cgaaaagtgc |      |
|            |            |            |            |             |            |      |

cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7920 cgaggccctt tcgtcttcaa gaattgtt 7948

```
<210> 37
<211> 7963
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pAdApt containing HIV genes
<400> 37
ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattqaaqc caatatgata atqaqqqqqt qqaqtttqtq acqtqqcqcq qqqcqtqqqa 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360 .
actcatageg egtaatattt gtetagggee geggggaett tgacegttta egtggagaet 420
cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480
tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840
catatgccaa gtacgcccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 960
gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020
tcacggggat ttccaagtct ccaccccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140
aggegtgtac ggtgggaggt ctatataagc agagetegtt tagtgaaccg teagategee 1200
tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260
cgtcaccgtc gtcgacacgt gtgatcagat atcgcggccg ctctagaatg cgcgtgaagg 1320
agaagtacca geacetgtgg egetgggget ggegetgggg caccatgetg etgggcatge 1380
tgatgatetg cagegecace gagaagetgt gggtgacegt gtactaegge gtgeeegtgt 1440
ggaaggaggc caccaccacc etgetetgeg ceagegaege caaggeetae gacacegagg 1500
tgcacaacgt gtgggccacc cacgcctgcg tgcccaccga ccccaacccc caggaggtgg 1560
tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcgtgaag ctgaccccc 1680
tgtgcgtgag cctgaagtgc accgacctga agaacgacac caacaccaac agcagcagcg 1740
gccgcatgat catggagaag ggcgagatca agaactgcag ettcaacatc agcaccagca 1800
teegeggeaa ggtgeagaag gagtaegeet tettetacaa getggaeate ateeceateg 1860
acaacgacac caccagctac agcetgacca getgeaacac cagcqtgatc acceaggeet 1920
gccccaaggt gagettegag cccatececa accactactg egccceegec ggettegeca 1980
tectgaagtg caaggacaag aagtteaacg geaagggeee etgeaceaac gtgageaceg 2040
tgcagtgcac ccaeggcate egeceegtgg tgagcacea getgetgetg aaeggtagee 2100
tggccgagga ggaggtggtg atccgcagcg ctaacttcgc cgacaacgcc aaggtgatca 2160
tegtgcaget gaacgagage gtggagatea actgcacceg egetagegee cactgcaace 2220
tgagccgcgc caagtggaac gacaccctga acaagatcgt gatcaagctg cgcgagcagt 2280
teggeaacaa gaccategtg tteaageaca geageggegg egacceegag ategtgacce 2340
acagetteaa etgeggegge gagttettet actgeaacag cacceagetg tteaacagea 2400
cctggttcaa cagcacctgg agcaccgagg gcagcaacaa caccqaggc agcqacacca 2460
```

|            | ctgccgcatc |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | ccccatcagc |            |            |            |            |      |
| tgacccgcga | cggcggcaac | agcaacaacg | agagcgagat | cttccgtccg | ggcggcggcg | 2640 |
| acatgcgcga | caactggcgc | agcgagctgt | acaagtacaa | ggtggtgaag | atcgagcccc | 2700 |
| tgggcgtggc | ccccaccaag | gccaagctta | ccgtccaggc | ccgccagctg | ctgagcggca | 2760 |
| tcgtgcagca | gcagaacaac | ctgctgcgcg | ccatcgaggc | ccagcagcac | ctgctgcagc | 2820 |
| tgaccgtgtg | gggcatcaag | cagctgcagg | cccgcaccct | ggccgtggag | cgctacctga | 2880 |
| aggaccagca | gcagctcgag | cagatctgga | accacaccac | ctggatggag | tgggaccgcg | 2940 |
|            | ctacaccagc |            |            |            |            |      |
|            | gcaggagctg |            |            |            |            |      |
|            | ctggctgtgg |            |            |            |            |      |
|            | cccctcccc  |            |            |            |            |      |
| tttcctaata | aaatgaggaa | attgcatcgc | attgtctgag | taggtgtcat | tctattctgg | 3240 |
|            | ggggcagcac |            |            |            |            |      |
|            | gggctctatg |            |            |            |            |      |
|            | agatctgtac |            |            |            |            |      |
|            | gtcttatgta |            |            |            |            |      |
|            | ttgatggaag |            |            |            |            |      |
|            | gtcagaatgt |            |            |            |            |      |
|            | ccttgaccta |            |            |            |            |      |
|            | cagccgctgc |            |            |            |            |      |
|            | caagcagtgc |            |            |            |            |      |
|            | aattggattc |            |            |            |            |      |
|            | agcaggtttc |            |            |            |            |      |
|            | aaccagactc |            |            |            |            |      |
|            | tgcgcgcgcg |            |            |            |            |      |
|            | ccaggacgtg |            |            |            |            |      |
|            | ggtggaggta |            |            |            |            |      |
|            | cgtagcagga |            |            |            |            |      |
|            | ggggcaggcc |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | gggatatgag |            |            |            |            |      |
|            | tccggggatt |            |            |            |            |      |
|            | tgtcatgtag |            |            |            |            |      |
|            | gattttccat |            |            |            |            |      |
|            | agatatttct |            |            |            |            |      |
|            | tttttacaaa |            |            |            |            |      |
|            | gggcgtagtt |            |            |            |            |      |
|            | tgtctacctg |            |            |            |            |      |
|            | aaagcaggtt |            |            |            |            |      |
|            | ttaccggctg |            |            |            |            |      |
|            | ccacttcgtt |            |            |            |            |      |
|            | gctcgccgcc |            |            |            |            |      |
|            | cgtccgccgt |            |            |            |            |      |
|            | cggtcacctg |            |            |            |            |      |
|            | gctttcgctg |            |            |            |            |      |
|            | cgggcgcagg |            |            |            |            |      |
|            | cgcgctggcc |            |            |            |            |      |
|            | gccctgcgcg |            |            |            |            |      |
|            | gtggcccttg |            |            |            |            |      |
|            | tttgagggcg |            |            |            |            |      |
|            | gcaggccccg |            |            |            |            |      |
| gttcggggtc | aaaaaccagg | tttcccccat | gctttttgat | gcgtttctta | cctctggttt | 5580 |
|            | gtgtccacgc |            |            |            |            |      |
|            | gtcctcgagc |            |            |            |            |      |
| ctgagacaaa | ggctcgcgtc | caggccagca | cgaaggaggc | taagtgggag | gggtagcggt | 5760 |

```
cgttgtccac tagggggtcc actcgctcca gggtgtgaag acacatgtcg ccctcttcgg 5820
catcaaggaa ggtgattggt ttgtaggtgt aggccacgtg accgggtgtt cctgaagggg 5880
ggctataaaa gggggtgggg gcgcgttcgt cctcactctc ttccgcatcg ctgtctgcga 5940
gggccagctg ttggggtgag tcgacgcgag gctggatggc cttccccatt atgattcttc 6000
tegetteegg eggeateggg atgeeegegt tgeaggeeat getgteeagg eaggtagatg 6060
acgaccatca gggacagctt caaggccagc aaaaggccag gaaccgtaaa aaggccgcgt 6120
tgctggcgtt tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa 6180
gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct 6240
ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc 6300
ettegggaag egtggegett teteataget eaegetgtag gtateteagt teggtgtagg 6360
tegttegete caagetggge tgtgtgeaeg aacceeegt teageeegae egetgegeet 6420
tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag 6480
cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga 6540
agtggtggcc taactacggc tacactagaa ggacagtatt tggtatctgc gctctgctga 6600
agccagttac etteggaaaa agagttggta getettgate eggeaaacaa accaeegetg 6660
gtagcggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag 6720
aagateettt gatettitet aeggggtetg aegeteagtg gaacgaaaac teaegttaag 6780
ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat 6840
gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct 6900
taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata gttgcctgac 6960
teccegtegt gtagataact acgatacggg agggettace atetggeece agtgetgeaa 7020
tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac cagccagccg 7080
gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag tctattaatt 7140
gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac gttgttgcca 7200
ttgctgcagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc agctccggtt 7260
cccaacgate aaggegagtt acatgatece ccatgttgtg caaaaaaggeg gttageteet 7320
teggteetee gategttgte agaagtaagt tggeegeagt gttateacte atggttatgg 7380
cagcactgca taattotett actgtcatgc catccgtaag atgcttttct gtgactggtq 7440
agtactcaac caagtcatto tgagaatagt gtatgcggcg accgagttgc tottgcccgg 7500
cgtcaacacg ggataatacc gcgccacata gcagaacttt aaaagtgctc atcattggaa 7560
aacgttette ggggegaaaa eteteaagga tettaceget gttgagatee agttegatgt 7620
aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc gtttctgggt 7680
gagcaaaaac aggaaggcaa aatgccgcaa aaaaggggaat aagggcgaca cggaaatgtt 7740
gaatactcat actetteett ttteaatatt attgaageat ttateagggt tattgtetea 7800
tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat 7860
ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca ttaacctata 7920
aaaataggcg tatcacgagg ccctttcgtc ttcaagaatt gtt
                                                                  7963
```

```
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pAdApt containing HIV genes

<400> 38
 ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
 ggattgaagc caatatgata atgaggggt ggagtttgtg acgtggcgg gggcgtggga 120
 acggggcggg tgacgtagta gtgtggcga agtgtgatgt tgcaagtgtg gcggaacaca 180
 tgtaagcgac ggatgtgca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240
 gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
 tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
```

<210> 38 <211> 7966

| acteatageg | cgtaatattt  | atchaggee  | acadagaett | taaccattta | catacaaact | 420  |
|------------|-------------|------------|------------|------------|------------|------|
|            | tttttctcag  |            |            |            |            |      |
|            |             |            |            |            |            |      |
|            | acgtaccagt  |            |            |            |            |      |
|            | tgtacattta. |            |            |            |            |      |
|            | gttattaata  |            |            |            |            |      |
|            | ttacataact  |            |            |            |            |      |
|            | cgtcaataat  |            |            |            |            |      |
|            | gggtggagta  |            |            |            |            |      |
|            | gtacgccccc  |            |            |            |            |      |
|            | tgaccttatg  |            |            |            |            |      |
|            | tggtgatgcg  |            |            |            |            |      |
|            | ttccaagtct  |            |            |            |            |      |
|            | actttccaaa  |            |            |            |            |      |
| aggcgtgtac | ggtgggaggt  | ctatataagc | agagctcgtt | tagtgaaccg | tcagatcgcc | 1200 |
| tggagacgcc | atccacgctg  | ttttgacctc | catagaagac | accgggaccg | atccagcctc | 1260 |
| cgtcaccgtc | gtcgacacgt  | gtgatcagat | atcgcggccg | ctctagaatg | cgcgtgaagg | 1320 |
|            | gcacctgtgg  |            |            |            |            |      |
|            | cagcgccacc  |            |            |            |            |      |
|            | caccaccacc  |            |            |            |            |      |
|            | gtgggccacc  |            |            |            |            |      |
|            | cgtgaccgag  |            |            |            |            |      |
|            | catcagcctg  |            |            |            |            |      |
|            | cctgaagtgc  |            |            |            |            |      |
|            | catggagaag  |            |            |            |            |      |
| tccgcggcaa | ggtgcagaag  | gagtacgcct | tcttctacaa | gctggacatc | atccccatcg | 1860 |
|            | caccagctac  |            |            |            |            |      |
|            | gagcttcgag  |            |            |            |            |      |
|            | caaggacaag  |            |            |            |            |      |
|            | ccacggcatc  |            |            |            |            |      |
|            | ggaggtggtg  |            |            |            |            |      |
|            | gaacgagagc  |            |            |            |            |      |
|            | catcggcccc  |            |            |            |            |      |
|            | ccactgcaac  |            |            |            |            |      |
|            | gcgcgagcag  |            |            |            |            |      |
|            | gatcgtgacc  |            |            |            |            |      |
|            | gccctgccgc  |            |            |            |            |      |
|            | ccccccatc   |            |            |            |            |      |
|            | cgacggcggc  |            |            |            |            |      |
|            | cgacaactgg  |            |            |            |            |      |
|            | ggcccccacc  |            |            |            |            |      |
|            | gcagcagaac  |            |            |            |            |      |
|            | gtggggcatc  |            |            |            |            |      |
|            | gcagcagete  |            |            |            |            |      |
|            | caactacacc  |            |            |            |            |      |
|            | cgagcaggag  |            |            |            |            |      |
|            | caactggctg  |            |            |            |            |      |
|            | ttgcccctcc  |            |            |            |            |      |
|            | ataaaatgag  |            |            |            |            |      |
|            | ggtggggcag  |            |            |            |            |      |
|            | ggtgggctct  |            |            |            |            |      |
|            | ataagatctg  |            |            |            |            |      |
|            | ggggtcttat  |            |            |            |            |      |
|            | cgtttgatgg  |            |            |            |            |      |
|            | tgcgtcagaa  |            |            |            |            |      |
|            | ctaccttgac  |            |            |            |            |      |
| -          | <b>-</b>    |            |            |            |            |      |

|            |            |            |            |            |            | 2000 |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | ctttgctttc |      |
|            |            |            |            |            | caagttgacg |      |
| gctcttttgg | cacaattgga | ttctttgacc | cgggaactta | atgtcgtttc | tcagcagctg | 3840 |
| ttggatctgc | gccagcaggt | ttctgccctg | aaggcttcct | cccctcccaa | tgcggtttaa | 3900 |
| aacataaata | aaaaaccaga | ctctgtttgg | atttggatca | agcaagtgtc | ttgctgtctt | 3960 |
|            |            |            |            |            | gttgagggtc |      |
|            |            |            |            |            | catgggcata |      |
|            |            |            |            |            | ggtggtgttg |      |
|            |            |            |            |            | tttcagtagc |      |
|            |            |            |            |            | aagctgggat |      |
|            |            |            |            |            | ggctatgttc |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | gtatccggtg |      |
|            |            |            |            |            | ggagacgccc |      |
|            |            |            |            |            | cccacgggcg |      |
|            |            |            |            |            | caggatgaga |      |
|            |            |            |            |            | tataatggtt |      |
|            |            |            |            |            | tttgagttca |      |
|            |            |            |            |            | ggtagggag  |      |
|            |            |            |            |            | ggtgggcccg |      |
| taaatcacac | ctattaccgg | ctgcaactgg | tagttaagag | agctgcagct | gccgtcatcc | 4860 |
| ctgagcaggg | gggccacttc | gttaagcatg | tccctgactc | gcatgttttc | cctgaccaaa | 4920 |
| tccgccagaa | ggcgctcgcc | gcccagcgat | agcagttctt | gcaaggaagc | aaagtttttc | 4980 |
| aacggtttga | gaccgtccgc | cgtaggcatg | cttttgagcg | tttgaccaag | cagttccagg | 5040 |
|            |            |            |            |            | tcctcgtttc |      |
|            |            |            |            |            | cgggccaggg |      |
|            |            |            |            |            | gtgaaggggt |      |
|            |            |            |            |            | gtgctgaagc |      |
|            |            |            |            |            | tcatagtcca |      |
|            |            |            |            |            | ccgcacgagg |      |
|            |            |            |            |            | tccggggagt |      |
|            |            |            |            |            | gtgagctctg |      |
|            |            |            |            |            | ttacctctgg |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | ccgtatacag |      |
|            |            |            |            |            | aactcggacc |      |
|            |            |            |            |            | gaggggtagc |      |
|            |            |            |            |            | tegecetett |      |
|            |            |            |            |            | gttcctgaag |      |
|            |            |            |            |            | tegetgtetg |      |
|            |            |            |            |            | attatgattc |      |
|            |            |            |            |            | aggcaggtag |      |
| atgacgacca | tcagggacag | cttcaaggcc | agcaaaaggc | caggaaccgt | aaaaaggccg | 6120 |
| cgttgctggc | gtttttccat | aggctccgcc | cccctgacga | gcatcacaaa | aatcgacgct | 6180 |
| caagtcagag | gtggcgaaac | ccgacaggac | tataaagata | ccaggcgttt | ccccctggaa | 6240 |
| gctccctcgt | gcgctctcct | gttccgaccc | tgccgcttac | cggatacctg | teegeettte | 6300 |
| tecetteggg | aagcgtggcg | ctttctcata | gctcacgctg | taggtatctc | agttcggtgt | 6360 |
|            |            |            |            |            | gaccgctgcg |      |
|            |            |            |            |            | tcgccactgg |      |
|            |            |            |            |            | acagagttct |      |
|            |            |            |            |            | tgcgctctgc |      |
|            |            |            |            |            | caaaccaccg |      |
|            |            |            |            |            | aaaggatete |      |
|            |            |            |            |            | aactcacgtt |      |
|            |            |            |            |            | ttaaattaaa |      |
|            |            |            |            |            | agttaccaat |      |
|            |            |            |            |            | atagttgcct |      |
| guitaattag | Lyayycacci | agega      | cocycocacc | cogucatou  | acagetyeet | 5500 |

```
gactccccgt cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg 7020
caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag 7080
ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta 7140
attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg 7200
ccattgctgc aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg 7260
gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct 7320
ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta 7380
tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg 7440
gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc 7500
eggegteaac acgggataat accgegeeac atageagaac tttaaaagtg ctcateattg 7560
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga 7620
tgtaacccac tcgtgcaccc aactgatett cagcatettt tactttcacc agcgtttctg 7680
ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat 7740
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 7800
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca 7860
catttecceg aaaagtgeea eetgaegtet aagaaaecat tattateatg acattaaect 7920
ataaaaatag gcgtatcacg aggccctttc qtcttcaaga attqtt
```

<210> 39 <211> 7870 <212> DNA <213> Artificial Sequence <220> <223> plasmid pAdApt containing HIV genes <400> 39 ttaattaacc gcaattetea tgtttgacag ettateatea teaataatat acettatttt 60 ggattgaagc caatatgata atgaggggt ggagtttgtg acgtggcgcg gggcgtggga 120 acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180 tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240 gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300 tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360 actcatageg egtaatattt gtetagggee geggggaett tgacegttta egtggagaet 420 cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480 tatagtcagt acgtaccagt gcactggcct agageggccc cattgcatac gttgtatcca 540 tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600 ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660 gagtteegeg ttacataact tacggtaaat ggeeegeetg getgacegee caacgacece 720 cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780 tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840 catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900 gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 960 gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020 tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080 aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140 aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200 tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260 cgtcaccgtc gtcgacacgt gtgatcagat atcgcggccg ctctagaatg cgcgtgaagg 1320 agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgqqcatqc 1380 tgatgatctg cagcgccacc gagaagctgt gggtgaccgt gtactacggc gtgcccgtgt 1440 ggaaggagge caccaccacc etgetetgeg ccagegaege caaggeetac gacaccgagg 1500 tgcacaacgt gtgggccacc cacgcctgcg tgcccaccga ccccaacccc caggaggtgg 1560

|            |            |            |            | cgacatggtg |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            | ctgcgtgaag |            |      |
|            |            |            |            | caacaccagc |            |      |
| aggcctgccc | caaggtgagc | ttcgagccca | tececaacea | ctactgcgcc | cccgccggct | 1800 |
| tcgccatcct | gaagtgcaag | gacaagaagt | tcaacggcaa | gggcccctgc | accaacgtga | 1860 |
| gcaccgtgca | gtgcacccac | ggcatccgcc | ccgtggtgag | cacccagctg | ctgctgaacg | 1920 |
|            |            |            |            | cttcgccgac |            |      |
|            |            |            |            | cacccgcccc |            |      |
|            |            |            |            | caccaccggc |            |      |
|            |            |            |            | gtggaacgac |            |      |
|            |            |            |            | catcgtgttc |            |      |
|            |            |            |            | cggcggcgag |            |      |
|            |            |            |            |            |            |      |
|            |            |            |            | cacctggagc |            |      |
|            |            |            |            | ccgcatcaag |            |      |
|            |            |            |            | catcagcggc |            |      |
|            |            |            |            | cggcaacagc |            |      |
|            |            |            |            | ctggcgcagc |            |      |
|            |            |            |            | caccaaggcc |            |      |
|            |            |            |            | gaacaacctg |            |      |
| tcgaggccca | gcagcacctg | ctgcagctga | ccgtgtgggg | catcaagcag | ctgcaggccc | 2760 |
| gcaccctggc | cgtggagcgc | tacctgaagg | accagcagca | gctcgagcag | atctggaacc | 2820 |
| acaccacctg | gatggagtgg | gaccgcgaga | tcaacaacta | caccagectg | atccacagcc | 2880 |
|            |            |            |            | ggagctgctg |            |      |
|            |            |            |            | gctgtggtga |            |      |
|            |            |            |            | ctccccgtg  |            |      |
|            |            |            |            | tgaggaaatt |            |      |
| atctgagtag | gtgtcattct | attetgagga | ataaaataaa | gcagcacagc | aaaaaaaaaa | 3180 |
| attoggaaga | caatagcagg | catactagag | atacaataaa | ctctatgggt | acccadacc  | 3240 |
| gcataacttc | gtataatgta | tactatacaa | acttataaca | tctgtactga | accougggee | 3300 |
|            |            |            |            | ttatgtagtt |            |      |
|            |            |            |            | atggaagcat |            |      |
|            |            |            |            |            |            |      |
| attestanta | cgcgcatgcc | cccatgggcc | ggggtgcgtc | agaatgtgat | gggeteeage | 3480 |
| actgatggte | geeeegreer | geeegeaaae | tetactacet | tgacctacga | gacegtgtet | 3540 |
|            |            |            |            | ccgctgcagc |            |      |
|            |            |            |            | gcagtgcagc |            |      |
|            |            |            |            | tggattcttt |            |      |
| cttaatgtcg | tttctcagca | gctgttggat | ctgcgccagc | aggtttctgc | cctgaaggct | 3780 |
|            |            |            |            | cagactctgt |            |      |
| atcaagcaag | tgtcttgctg | tctttattta | ggggttttgc | gcgcgcggta | ggcccgggac | 3900 |
|            |            |            |            | ggacgtggta |            |      |
| tggatgttca | gatacatggg | cataagcccg | tctctggggt | ggaggtagca | ccactgcaga | 4020 |
| gcttcatgct | gcggggtggt | gttgtagatg | atccagtcgt | agcaggagcg | ctgggcgtgg | 4080 |
|            |            |            |            | gcaggccctt |            |      |
|            |            |            |            | atatgagatg |            |      |
|            |            |            |            | ggggattcat |            |      |
|            |            |            |            | catgtagctt |            |      |
| acatagaaga | acttggagac | accettataa | cctccaagat | tttccatgca | ttcatccata | 4380 |
|            |            |            |            | tatttctggg |            |      |
|            |            |            |            | ttacaaagcg |            |      |
|            |            |            |            | cgtagttacc |            |      |
|            |            |            |            |            |            |      |
|            |            |            |            | ctacctgcgg |            |      |
| aaaacygttt | ccyyggtagg | ggagaccage | cyyyaayaaa | gcaggttcct | yagcagctgc | 4680 |
|            |            |            |            | ccggctgcaa |            |      |
| agagagccgc | agetgeegte | acccctgage | aggggggcca | cttcgttaag | cargreectg | 4800 |
| actcgcatgt | tttccctgac | caaatccgcc | agaaggcgct | cgccgcccag | cgatagcagt | 4860 |
|            |            |            |            |            |            | •    |

|              |            |              | 1.1         |            |                          | 4020      |
|--------------|------------|--------------|-------------|------------|--------------------------|-----------|
| tcttgcaagg   | aagcaaagtt | tttcaacggt   | ttgagaccgt  | ccgccgragy | catgcttttg               | 4000      |
| agcgtttgac   | caagcagttc | caggcggtcc   | cacagetegg  | tcacctgctc | tacggcatct               | 4980      |
| cgatccagca   | tatctcctcg | tttcgcgggt   | tggggcggct  | ttcgctgtac | ggcagtagtc               | 5040      |
| ggtgctcgtc   | cagacgggcc | agggtcatgt   | ctttccacgg  | gcgcagggtc | ctcgtcagcg               | 5100      |
| tagtctgggt   | cacggtgaag | gggtgcgctc   | cgggctgcgc  | gctggccagg | gtgcgcttga               | 5160      |
| ggctggtcct   | gctggtgctg | aagcgctgcc   | ggtcttcgcc  | ctgcgcgtcg | gccaggtagc               | 5220      |
| atttgaccat   | ggtgtcatag | tccagcccct   | ccgcggcgtg  | gcccttggcg | cgcagcttgc               | 5280      |
| ccttggagga   | ggcgccgcac | gaggggcagt   | gcagactttt  | gagggcgtag | agcttgggcg               | 5340      |
| cgagaaatac   | cgattccggg | gagtaggcat   | ccgcgccgca  | ggccccgcag | acggtctcgc               | 5400      |
| attccacgag   | ccaggtgagc | tctggccgtt   | cggggtcaaa  | aaccaggttt | ccccatgct                | 5460      |
| ttttgatgcg   | tttcttacct | ctggtttcca   | tgagccggtg  | tccacgctcg | gtgacgaaaa               | 5520      |
| ggctgtccgt   | gtccccgtat | acagacttga   | gaggcctgtc  | ctcgagcggt | gttccgcggt               | 5580      |
| cctcctcgta   | tagaaactcg | gaccactctg   | agacaaaggc  | tcgcgtccag | gccagcacga               | 5640      |
| aggaggetaa   | ataggaggg  | tagcggtcgt   | tgtccactag  | ggggtccact | cgctccaggg               | 5700      |
| totoaagaca   | catqtcqccc | tetteggeat   | caaggaaggt  | gattggtttg | taggtgtagg               | 5760      |
| ccacqtqacc   | aggtattcct | gaaggggggc   | tataaaaggg  | ggtgggggg  | cgttcgtcct               | 5820      |
| cactctcttc   | cgcatcgctg | tctgcgaggg   | ccagctgttg  | gggtgagtcg | acgcgaggct               | 5880      |
| ggatggcctt   | ccccattatq | attetteteq   | cttccggcgg  | catcgggatg | cccgcgttgc               | 5940      |
| aggccatgct   | gtccaggcag | gtagatgacg   | accatcaggg  | acagetteaa | ggccagcaaa               | 6000      |
| aggccaggaa   | ccgtaaaaag | accacattac   | tggcgttttt  | ccataggete | cgccccctg                | 6060      |
| acgagcatca   | caaaaatcga | cactcaaatc   | agaggtggcg  | aaacccgaca | ggactataaa               | 6120      |
| gataccagge   | atttcccct  | ggaageteee   | tcatacactc  | tcctqttccq | accetgeege               | 6180      |
| ttaccocata   | cetatecace | tttctccctt   | caddaadcat  | ggcgctttct | catagctcac               | 6240      |
| actotagata   | teteagtteg | atatagatca   | ttcqctccaa  | gctgggctgt | gtgcacgaac               | 6300      |
| ccccattca    | acceasees  | tacaccttat   | ccaataacta  | tcatcttgag | tccaacccgg               | 6360      |
| taanacacna   | cttatcacca | chaacaacaa   | ccactggtaa  | caggattagc | agagcgaggt               | 6420      |
| atatagacacga | tactacagaa | ttcttgaagt   | gataacctaa  | ctacqqctac | actagaagga               | 6480      |
| cagtatttcc   | tatctgcgct | ctactaaaac   | cagttacctt  | cqqaaaaaqa | gttggtagct               | 6540      |
| cttcatccgg   | caeeegegee | accoctoota   | acaataattt  | ttttatttac | aagcagcaga               | 6600      |
| ttacacacaa   | 22222222   | totcaagaag   | atcctttgat  | cttttctacg | gggtctgacg               | 6660      |
| ctaegegeag   | cassactca  | cattaaggga   | htttggtgat  | gagattatca | aaaaggatct               | 6720      |
| tcacctagat   | ccttttaaat | taaaaatgaa   | gttttaaatc  | aatctaaagt | atatatgagt               | 6780      |
| aaacttaatc   | taacaattac | caatoottaa   | tcagtgaggc  | acctatctca | gcgatctgtc               | 6840      |
| tatttcattc   | atccatactt | acctgactcc   | ccatcatata  | gataactacg | atacgggagg               | 6900      |
| cattecgete   | tagacacaat | actacaataa   | taccocoaga  | cccacactca | ccggctccag               | 6960      |
| atttatcacc   | astasaccad | ccadcadaa    | adaccasaca  | cagaagtggt | cctgcaactt               | 7020      |
| tataggata    | catecagtet | attaattatt   | accadasac   | tagagtaagt | agttcgccag               | 7080      |
| taceegeete   | gegeaacgtt | attaccatta   | ctgcaggaage | cataatatca | cgctcgtcgt               | 7140      |
| ttaatayee    | ttcattcacc | tecaattece   | aacqatcaaq  | acaaattaca | tgatccccca               | 7200      |
| tattatass    | aaaaaaaatt | agetectted   | atcetecat   | cattatcaga | agtaagttgg               | 7260      |
| egetgtgtaa   | addageggee | attatageag   | cactocataa  | ttctcttact | gtcatgccat               | 7320      |
| cegeagegee   | attactatg  | actactacat   | actraaccaa  | atcattctaa | gaatagtgta               | 7380      |
| cegtaagatg   | anattastat | taccagaga    | caacaccaa   | taataccoco | ccacatagca               | 7440      |
| tgeggegaee   | gaguiguicu | attogaaaa    | ottettegga  | acassastc  | tcaaggatct               | 7500      |
| gaactttaaa   | agratecact | togatataac   | ccactcatac  | acccaactca | tcttcagcat               | 7560      |
| tacegetgtt   | gagacccagc | togatgtaac   | caacacgaga  | acccaaccga | accacagaa                | 7620      |
| CCCCCacccc   | caccagegee | cougggugag   | tactcatact  | cttcctttt  | gccgcaaaaa<br>caatattatt | 7680      |
| agggaataag   | ggegaeaegg | adatyrtydd   | cacccacact  | atttgaatgt | atttagaaaa               | 7740      |
| gaagcattta   | Leagggttat | . cgccccacga | grayaracat  | accegaacge | atchagaaaa               | 7800      |
| acaaacaaat   | aggggttccg | regeacatite  | ataggggatat | gecaeeegae | gtctaagaaa               | 7860      |
|              |            | accididada   | acayycycat  | cacyayyccc | tttcgtcttc               | 7870      |
| aagaattgtt   |            |              |             |            |                          | , , , , , |
|              |            |              |             |            |                          |           |

<210> 40

<211> 7906 <212> DNA <213> Artificial Sequence <220> <223> plasmid pAdApt containing HIV genes <400> 40 ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60 ggattgaagc caatatgata atgaggggt ggagtttgtg acgtggcgcg gggcgtggga 120 acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180 tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240 gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300 tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360 actcatageg egtaatattt gtetagggee geggggaett tgacegttta egtggagaet 420 cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaaqttqqc qttttattat 480 tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540 tatcataata tgtacattta tattggctca tgtccaacat taccqccatq ttqacattqa 600 ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660 gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720 cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780 tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840 catatgccaa gtacgcccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900 gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 960 gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020 tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080 aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140 aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200 tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260 eqteacegte gtegacacgt gtgateagat ategeggeeg etetagaatg egegtgaagg 1320 agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380 tgatgatetg cagegecace gagaagetgt gggtgacegt gtactaegge gtgecegtgt 1440 ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500 tgcacaacgt gtgggccacc cacgcctgcg tgcccaccga ccccaacccc caggaggtgg 1560 tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620 acgaggacat catcagcetg tgggaccaga gcctgaagcc ctgcgtgaag ctgaccccc 1680 tgtgcgtgag cctgaagtgc accgacgcta gcaagaactg cagcttcaac atcagcacca 1740 gcatccgcgg caaggtgcag aaggagtacg ccttcttcta caagctggac atcatcccca 1800 tegacaacga caccaccage tacageetga ecagetgeaa caccagegtg ateacccagg 1860 cetgececaa ggtgagette gageceatee ecaaceaeta etgegeecee geeggetteg 1920 ccatcctgaa gtgcaaggac aagaagttca acggcaaggg cccctgcacc aacgtgagca 1980 ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg ctgaacggta 2040 gcctggccga ggaggaggtg gtgatccgca gcgctaactt cgccgacaac gccaaqqtga 2100 tcatcgtgca gctgaacgag agcgtggaga tcaactgcac ccgcgctagc gcccactgca 2160 acctgagccg cgccaagtgg aacgacaccc tgaacaagat cgtgatcaag ctgcgcgagc 2220 agttcggcaa caagaccatc gtgttcaagc acagcagcgg cggcgacccc gagatcgtga 2280

cccacagett caactgegge ggegagttet tetactgeaa cageacccag etgttcaaca 2340 gcacctggtt caacagcacc tggagcaccg agggcagcaa caacaccgag ggcagcgaca 2400 ccatcaccct gccctgccgc atcaagcaga tcatcaacat gtggcagaag gtgggcaagg 2460 ccatgtacgc ccccccatc ageggccaga tecgetgcag cagcaacatc aceggcetgc 2520 tgctgacccg cgacggcggc aacagcaaca acgagagcga gatcttccgt ccgggcggcg 2580 gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg aagatcgagc 2640 ccctgggcgt ggcccccacc aaggccaagc ttaccgtcca ggcccgccag ctgctgagcg 2700 gcatcgtgca gcagcagaac aacctgctgc gcgccatcga ggcccagcag cacctgctgc 2760 agetgacegt gtggggcate aageagetge aggecegeae eetggeegtg gagegetace 2820

| tgaaggacca   | acaacaactc   | gagcagatct | ggaaccacac   | cacctggatg   | gagtgggacc | 2880 |
|--------------|--------------|------------|--------------|--------------|------------|------|
| acaagatcaa   | caachacacc   | agectgatec | acagectgat   | cgaggagagc   | cagaaccagc | 2940 |
| accacaacaa   | cgagcaggag   | ctactagage | tagacaagta   | ggccagcctg   | tggaactggt | 3000 |
| tcaacatcac   | caactageta   | tagtaaggat | ccagatctgc   | tataccttct   | agttgccagc | 3060 |
| catctattat   | ttacccctcc   | cccatacctt | ccttgaccct   | qqaaqqtgcc   | actcccactg | 3120 |
| tcctttccta   | ataaaatgag   | gaaattgcat | cacattatet   | gagtaggtgt   | cattctattc | 3180 |
| tagaggatag   | aataaaaacaa  | cacagcaagg | gggaggattg   | qqaaqacaat   | agcaggcatg | 3240 |
| ctagagagatag | ggtggggtus   | ataggtaccc | agggcggcat   | aacttcqtat   | aatgtatgct | 3300 |
| ataccasact   | ataagatetg   | tactgaaatg | tataoacata   | acttaagggt   | gggaaagaat | 3360 |
| atatgaagte   | acaagacceg   | graghthtgt | atctatttta   | cagcagccgc   | cgccgccatg | 3420 |
| acacaaggcg   | catttaataa   | aagcattgtg | ageteatatt   | tgacaacgcg   | catgccccca | 3480 |
| tarassassass | tacatcacaa   | tataataaa  | tccagcattg   | ataatcaccc   | cgtcctgccc | 3540 |
| cgggccgggg   | chacettese   | chacqaqacc | gtgtctggaa   | caccattaga   | gactgcagcc | 3600 |
| geaaacceca   | ottoragas    | tacaaccacc | acccccaaaa   | trataactaa   | ctttgctttc | 3660 |
| ceegeegeeg   | ttaasaaaa    | tacaacttcc | cattcatcca   | cccacaataa   | caagttgacg | 3720 |
| ctgageeege   | cigcaagcag   | thatttaaca | cagaactta    | atotoottto   | tcagcagctg | 3780 |
| getettttgg   | cacaactgga   | thatagata  | agggaacca    | ccctccaa     | tgcggtttaa | 3840 |
| ttggatctgc   | gccagcaggt   | etetgeeetg | atttaatta    | accepteda    | ttactatett | 3900 |
| aacataaata   | aaaaaccaga   | eccepting  | actiggatea   | ageaagegee   | ttgctgtctt | 3960 |
| tatttagggg   | ttttgcgcgc   | geggtaggee | egggaccagc   | battangata   | gttgagggtc | 4020 |
| ctgtgtattt   | tttccaggac   | gtggtaaagg | tgactetgga   | cyttcagata   | catgggcata | 4020 |
| agcccgtctc   | tggggtggag   | gtagcaccac | Egcagagett   | tacatata     | ggtggtgttg | 4140 |
| tagatgatcc   | agtcgtagca   | gagcaccaa  | gegragee     | caaaaacgcc   | tttcagtagc | 4200 |
| aagctgattg   | ccaggggcag   | gcccttggtg | taagtgttta   | caaageggee   | aagctgggat | 4260 |
| gggtgcatac   | gtggggatat   | gagatgcatc | ttggactgta   | tttttaggtt   | ggctatgttc | 4200 |
| ccagccatat   | ccctccgggg   | attcatgttg | tgcagaacca   | ccagcacagt   | gtatccggtg | 4320 |
| cacttgggaa   | atttgtcatg   | tagcttagaa | ggaaatgcgt   | ggaagaactt   | ggagacgccc | 4380 |
| ttgtgacctc   | caagattttc   | catgcattcg | tccataatga   | tggcaatggg   | cccacgggcg | 4440 |
| gcggcctggg   | cgaagatatt   | tctgggatca | ctaacgtcat   | agttgtgttc   | caggatgaga | 4500 |
| tcgtcatagg   | ccatttttac   | aaagcgcggg | cggagggtgc   | cagactgcgg   | tataatggtt | 4560 |
| ccatccggcc   | caggggcgta   | gttaccctca | cagatttgca   | tttcccacgc   | tttgagttca | 4620 |
| gatggggga    | tcatgtctac   | ctgcggggcg | atgaagaaaa   | cggtttccgg   | ggtaggggag | 4680 |
| atcagctggg   | aagaaagcag   | gttcctgagc | agetgegact   | taccgcagcc   | ggtgggcccg | 4740 |
| taaatcacac   | ctattaccgg   | ctgcaactgg | tagttaagag   | agetgeaget   | gccgtcatcc | 4800 |
| ctgagcaggg   | gggccacttc   | gttaagcatg | tecetgacte   | gcatgttttc   | cctgaccaaa | 4860 |
| tccgccagaa   | ggcgctcgcc   | gcccagcgat | agcagttctt   | gcaaggaagc   | aaagtttttc | 4920 |
| aacggtttga   | gaccgtccgc   | cgtaggcatg | cttttgagcg   | tttgaccaag   | cagttccagg | 4980 |
| cggtcccaca   | gctcggtcac   | ctgctctacg | gcatctcgat   | ccagcatatc   | tectegttte | 5040 |
| gcgggttggg   | gcggctttcg   | ctgtacggca | gtagtcggtg   | ctcgtccaga   | cgggccaggg | 5100 |
| tcatgtcttt   | ccacgggcgc   | agggtcctcg | tcagcgtagt   | ctgggtcacg   | gtgaaggggt | 5160 |
| gcgctccggg   | ctgcgcgctg   | gccagggtgc | gcttgaggct   | ggtcctgctg   | gtgctgaagc | 5220 |
| gctgccggtc   | ttcgccctgc   | gcgtcggcca | ggtagcattt   | gaccatggtg   | tcatagtcca | 5280 |
| gcccctccgc   | ggcgtggccc   | ttggcgcgca | gcttgccctt   | ggaggaggcg   | ccgcacgagg | 5340 |
| ggcagtgcag   | acttttgagg   | gcgtagagct | tgggcgcgag   | aaataccgat   | tccggggagt | 5400 |
| aggcatccgc   | gccgcaggcc   | ccgcagacgg | tctcgcattc   | cacgagccag   | gtgagctctg | 5460 |
| gccgttcggg   | gtcaaaaacc   | aggtttcccc | catgcttttt   | gatgcgtttc   | ttacctctgg | 5520 |
| tttccatgag   | ccggtgtcca   | cgctcggtga | . cgaaaaggct | gtccgtgtcc   | ccgtatacag | 5580 |
| acttgagagg   | cctgtcctcg   | agcggtgttc | : cgcggtcctc | ctcgtataga   | aactcggacc | 5640 |
| actctgagac   | aaaggctcgc   | gtccaggcca | gcacgaagga   | ggctaagtgg   | gaggggtagc | 5700 |
| ggtcgttgtc   | cactaggggg   | tccactcgct | ccagggtgtg   | aagacacatg   | tegecetett | 5760 |
| cggcatcaag   | gaaggtgatt   | ggtttgtagg | tgtaggccac   | gtgaccgggt   | gttcctgaag | 5820 |
| gggggctata   | aaagggggtg   | ggggcgcgtt | . cgtcctcact | ctcttccgca   | tcgctgtctg | 5880 |
| cgagggccag   | ctgttggggt   | gagtcgacgc | : gaggctggat | ggccttcccc   | attatgattc | 5940 |
| ttctcgcttc   | : cggcggcatc | gggatgcccg | r cgttgcaggc | catgctgtcc   | aggcaggtag | 6000 |
| atgacgacca   | tcagggacag   | cttcaaggco | : agcaaaaggc | : caggaaccgt | aaaaaggccg | 6060 |
| cgttgctggc   | gtttttccat   | aggctccgcc | cccctgacga:  | gcatcacaaa   | aatcgacgct | 6120 |

```
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa 6180
getecetegt gegetetect gtteegaeee tgeegettae eggataeetg teegeettte 6240
tecetteggg aagegtggeg ettteteata geteaegetg taggtatete agtteggtgt 6300
agglegting circaagetg ggetgtgtge acgaaceee egiteageee gaceqetqeq 6360
cettateegg taactategt ettgagteea acceqqtaaq acacqaetta teqceactqq 6420
cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 6480
tgaagtggtg gcctaactac ggctacacta gaaggacagt atttggtatc tgcqctctqc 6540
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 6600
ctggtagcgg tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 6660
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 6720
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 6840
gettaateag tgaggeacet ateteagega tetgtetatt tegtteatee atagetgeet 6900
gactccccgt cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg 6960
caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag 7020
ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta 7080
attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttqttq 7140
ccattgctgc aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg 7200
gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcqqttaqct 7260
ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatqqtta 7320
tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg 7380
gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgaqt tqctcttqcc 7440
cggcgtcaac acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg 7500
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcqa 7560
tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 7620
ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat 7680
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 7740
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca 7800
catttccccg aaaagtgcca cctgacgtct aagaaaccat tattatcatg acattaacct 7860
ataaaaatag gcgtatcacg aggccctttc gtcttcaaga attgtt
```

```
<210> 41
<211> 7909
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pAdApt containing HIV genes
<400> 41
```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataata accttattt 60 ggattgaagc caatatgata atgaggggt ggagtttgtg acgtggcgcg gggcgtggga 120 acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180 tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240 gacaattttc gcgcggttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300 tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360 actcatagcg cgtaatattt gtctagggcc gcggggactt tgaccgtta cgtggagact 420 cgcccaggtg tttttccag gtgttttccg cgttccgggt caaagttggc gttttattat 480 tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgacattga 600 ttattgacta gttattaata gtaatcaatt acgggacatt taggtccatg ccaacagt cccaataatg 660 gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720 cgcccattga cgtcaataat ggcgtatgtt cccaatagtaa cgccaataag gactttccat 780

|            |            |            |            |            | tcaagtgtat |         |
|------------|------------|------------|------------|------------|------------|---------|
| catatgccaa | gtacgccccc | tattgacgtc | aatgacggta | aatggcccgc | ctggcattat | 900     |
| gcccagtaca | tgaccttatg | ggactttcct | acttggcagt | acatctacgt | attagtcatc | 960     |
| gctattacca | tggtgatgcg | gttttggcag | tacatcaatg | ggcgtggata | gcggtttgac | 1020    |
|            |            |            |            |            | ttggcaccaa |         |
|            |            |            |            |            | aatgggcggt |         |
|            |            |            |            |            | tcagatcgcc |         |
|            |            |            |            |            | atccagcctc |         |
|            |            |            |            |            | cgcgtgaagg |         |
|            |            |            |            |            | ctgggcatgc |         |
|            |            |            |            |            |            |         |
|            |            |            |            |            | gtgcccgtgt |         |
|            |            |            |            |            | gacaccgagg |         |
|            |            |            |            |            | caggaggtgg |         |
|            |            |            |            |            | gagcagatgc |         |
|            |            |            |            |            | ctgacccccc |         |
| tgtgcgtgag | cctgaagtgc | accgacgcta | gcaagaactg | cagcttcaac | atcagcacca | 1740    |
| gcatccgcgg | caaggtgcag | aaggagtacg | ccttcttcta | caagctggac | atcatcccca | 1800    |
| tcgacaacga | caccaccagc | tacagcctga | ccagctgcaa | caccagcgtg | atcacccagg | 1860    |
| cctgccccaa | ggtgagcttc | gagcccatcc | ccaaccacta | ctgcgccccc | geeggetteg | 1920    |
| ccatcctgaa | gtgcaaggac | aagaagttca | acggcaaggg | cccctgcacc | aacgtgagca | 1980    |
|            |            |            |            |            | ctgaacggta |         |
|            |            |            |            |            | gccaaggtga |         |
|            |            |            |            |            | aacaacaccc |         |
|            |            |            |            |            | atcatcggcg |         |
|            |            |            |            |            | ctgaacaaga |         |
|            |            |            |            |            | cacagcagcg |         |
|            |            |            |            |            | ttctactgca |         |
|            |            |            |            |            | aaggtgggca |         |
|            |            |            |            |            |            |         |
|            |            |            |            |            | atcaccggcc |         |
|            |            |            |            |            | cgtccgggcg |         |
|            |            |            |            |            | gtgaagatcg |         |
|            |            |            |            |            | cagctgctga |         |
|            |            |            |            |            | cagcacctgc |         |
|            |            |            |            |            | gtggagcgct |         |
|            |            |            |            |            | atggagtggg |         |
|            |            |            |            |            | agccagaacc |         |
|            |            |            |            |            | ctgtggaact |         |
| ggttcaacat | caccaactgg | ctgtggtgag | gatccagatc | tgctgtgcct | tctagttgcc | 3060    |
| agccatctgt | tgtttgcccc | tcccccgtgc | cttccttgac | cctggaaggt | gccactccca | 3120    |
|            |            |            |            |            | tgtcattcta |         |
| ttctgggggg | tggggtgggg | cagcacagca | agggggagga | ttgggaagac | aatagcaggc | 3240    |
| atgctgggga | tgcggtgggc | tctatgggta | cccagggccg | cataacttcg | tataatgtat | 3300    |
| gctatacgaa | gttataagat | ctgtactgaa | atgtgtgggc | gtggcttaag | ggtgggaaag | 3360    |
| aatatataag | gtgggggtct | tatgtagttt | tgtatctgtt | ttgcagcagc | cgccgccgcc | 3420    |
|            |            |            |            |            | gcgcatgccc |         |
| ccatgggccg | gggtgcgtca | gaatgtgatg | ggctccagca | ttgatggtcg | ccccgtcctg | 3540    |
|            |            |            |            |            | ggagactgca |         |
|            |            |            |            |            | tgactttgct |         |
|            |            |            |            |            | tgacaagttg |         |
|            |            |            |            |            | ttctcagcag |         |
|            |            |            |            |            | caatgcggtt |         |
|            |            |            |            |            | gtcttgctgt |         |
|            |            |            |            |            | gtcgttgagg |         |
|            |            |            |            |            | atacatgggc |         |
|            |            |            |            |            |            |         |
| acaageeege | cccggggcg  | gaggrageae | uuccycayay | crucatycig | cggggtggtg | 3 V O V |

|             |                      |            |            | gcctaaaaat |            |      |
|-------------|----------------------|------------|------------|------------|------------|------|
|             |                      |            |            | ttacaaagcg |            |      |
|             |                      |            |            | gtatttttag |            |      |
| ttcccagcca  | tatccctccg           | gggattcatg | ttgtgcagaa | ccaccagcac | agtgtatccg | 4320 |
| gtgcacttgg  | gaaatttgtc           | atgtagctta | gaaggaaatg | cgtggaagaa | cttggagacg | 4380 |
|             |                      |            |            | tgatggcaat |            |      |
|             |                      |            |            | catagttgtg |            |      |
|             |                      |            |            | tgccagactg |            |      |
| gttccatccg  | gcccaggggc           | gtagttaccc | tcacagattt | gcatttccca | cgctttgagt | 4620 |
| tcagatgggg  | ggatcatgtc           | tacctgcggg | gcgatgaaga | aaacggtttc | cggggtaggg | 4680 |
|             |                      |            |            | acttaccgca |            |      |
|             |                      |            |            | gagagctgca |            |      |
| tecetgagea  | ggggggccac           | ttcgttaagc | atgtccctga | ctcgcatgtt | ttccctgacc | 4860 |
|             |                      |            |            | cttgcaagga |            |      |
|             |                      |            |            | gcgtttgacc |            |      |
|             |                      |            |            | gatccagcat |            |      |
|             |                      |            |            | gtgctcgtcc |            |      |
| gggtcatgtc  | tttccacggg           | cgcagggtcc | tcgtcagcgt | agtctgggtc | acggtgaagg | 5160 |
|             |                      |            |            | gctggtcctg |            |      |
|             |                      |            |            | tttgaccatg |            |      |
| ccaqcccctc  | cacaacataa           | cccttqqcqc | qcagcttgcc | cttggaggag | gcgccgcacg | 5340 |
| aggggagtg   | cagacttttg           | agggcgtaga | gettgggege | gagaaatacc | gattccgggg | 5400 |
|             |                      |            |            | ttccacgagc |            |      |
| ctaaccattc  | ggggtcaaaa           | accaggtttc | ccccatqctt | tttgatgcgt | ttcttacctc | 5520 |
| tootttccat  | gagccggtgt           | ccacqctcqq | tgacgaaaag | gctgtccgtg | tccccgtata | 5580 |
|             |                      |            |            | ctcctcgtat |            |      |
|             |                      |            |            | ggaggctaag |            |      |
|             |                      |            |            | gtgaagacac |            |      |
|             |                      |            |            | cacgtgaccg |            |      |
|             |                      |            |            | actctcttcc |            |      |
|             |                      |            |            | gatggccttc |            |      |
|             |                      |            |            | ggccatgctg |            |      |
|             |                      |            |            | ggccaggaac |            |      |
| °ccgcgttgct |                      |            |            |            |            |      |
|             |                      |            |            | ataccaggcg |            |      |
|             |                      |            |            | taccggatac |            |      |
|             |                      |            |            | ctgtaggtat |            |      |
|             |                      |            |            | ccccgttcag |            |      |
|             |                      |            |            | aagacacgac |            |      |
|             |                      |            |            | tgtaggcggt |            |      |
|             |                      |            |            | agtatttggt |            |      |
|             |                      |            |            | ttgatccggc |            |      |
|             |                      |            |            | tacgcgcaga |            |      |
|             |                      |            |            | tcagtggaac |            |      |
|             |                      |            |            | cacctagatc |            |      |
| aaaaatgaag  | ttttaaatca           | atctaaaqta | tatatgagta | aacttggtct | gacagttacc | 6840 |
| aatgcttaat  | cagtgaggca           | cctatctcag | cgatctgtct | atttcgttca | tccatagttg | 6900 |
| cctgactccc  | cgtcgtgtag           | ataactacga | tacgggaggg | cttaccatct | ggccccagtq | 6960 |
| ctgcaatgat  | accgcgagac           | ccacgctcac | cggctccaga | tttatcagca | ataaaccagc | 7020 |
| cadccddaad  | ggccgagcgc           | agaagtqqtc | ctgcaacttt | atccgcctcc | atccagtcta | 7080 |
| ttaattatta  | ccgggaagct           | agagtaagta | gttcgccagt | taatagtttg | cgcaacgttg | 7140 |
|             |                      |            |            | tggtatggct |            |      |
|             |                      |            |            | gttgtgcaaa |            |      |
|             |                      |            |            | cgcagtgtta |            |      |
| ttatqqcaqc  | actgcataat           | tetettactq | tcatgccatc | cgtaagatgc | ttttctgtga | 7380 |
|             | <b>3</b> ··· - • • • |            | -          |            |            |      |

```
ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt 7440 gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 7500 ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 7560 cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt 7620 ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaaa gggaataaagg gcgacacgga 7680 aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat cagggttatt 7740 gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 7800 gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa 7860 cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattgtt
```

<210> 42
<211> 7867
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pAdApt containing HIV genes

<400> 42 ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60 ggattgaagc caatatgata atgaggggt ggagtttgtg acgtggcgcg gggcgtggga 120 acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180 tqtaaqcqac qqatqtgqca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240 gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300 tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360 actcatagcg cgtaatattt gtctagggcc gcggggactt tgaccgttta cgtggagact 420 cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480 tataqtcaqt acqtaccaqt gcactggcct agagcggccc cattgcatac gttgtatcca 540 tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600 ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660 gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720 cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780 tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840 catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900 gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 960 gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020 tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080 aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140 aggegtgtac ggtgggaggt ctatataagc agagetegtt tagtgaaccg tcagategec 1200 tggagacgec atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260 cgtcaccgtc gtcgacacgt gtgatcagat atcgcggccg ctctagaatg cgcgtgaagg 1320 agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380 tgatgatctg cagegecacc gagaagetgt gggtgaccgt gtactacgge gtgcccgtgt 1440 qqaaqqaqqc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500 tgcacaacgt gtgggccacc cacgcctgcg tgcccaccga ccccaacccc caggaggtgg 1560 tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620 acqaqqacat catcaqcctg tgggaccaga gcctgaagcc ctgcgtgaag ctgaccccc 1680 tgtgcgtgag cctgaagtgc accgacctga agaacgacac caacaccaac agcagcagcg 1740 gccgcatgat catggagaag ggcgagatca agaactgcag cttcgctagc accagctgca 1800 acaccaqcqt qatcacccaq gcctgcccca aggtgagctt cgagcccatc cccaaccact 1860 actgcgccc cgccggcttc gccatcctga agtgcaagga caagaagttc aacggcaagg 1920 gcccctgcac caacgtgagc accgtgcagt gcacccacgg catccgcccc gtggtgagca 1980 cccagctgct gctgaacggt agcctggccg aggaggaggt ggtgatccgc agcgctaact 2040

|            | _          |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            | agctgaacga |            |            |      |
| cccgcgctag | cgcccactgc | aacctgagcc | gcgccaagtg | gaacgacacc | ctgaacaaga | 2160 |
| tcgtgatcaa | gctgcgcgag | cagttcggca | acaagaccat | cgtgttcaag | cacagcagcg | 2220 |
| gcggcgaccc | cgagatcgtg | acccacagct | tcaactgcgg | cggcgagttc | ttctactgca | 2280 |
| acagcaccca | gctgttcaac | agcacctggt | tcaacagcac | ctggagcacc | gagggcagca | 2340 |
| acaacaccga | gggcagcgac | accatcaccc | tgccctgccg | catcaagcag | atcatcaaca | 2400 |
| tgtggcagaa | ggtgggcaag | gccatgtacg | cccccccat  | cagcggccag | atccgctgca | 2460 |
| gcagcaacat | caccggcctg | ctgctgaccc | gcgacggcgg | caacagcaac | aacgagagcg | 2520 |
|            |            |            | gcgacaactg |            |            |      |
| acaaggtggt | gaagatcgag | cccctgggcg | tggcccccac | caaggccaag | cttaccgtcc | 2640 |
| aggcccgcca | gctgctgagc | ggcatcgtgc | agcagcagaa | caacctgctg | cgcgccatcg | 2700 |
| aggcccagca | gcacctgctg | cagctgaccg | tgtggggcat | caagcagctg | caggcccgca | 2760 |
|            |            |            | agcagcagct |            |            |      |
|            |            |            | acaactacac |            |            |      |
|            |            |            | acgagcagga |            |            |      |
| gggccagcct | gtggaactgg | ttcaacatca | ccaactggct | gtggtgagga | tccagatctg | 3000 |
| ctgtgccttc | tagttgccag | ccatctgttg | tttgcccctc | ccccgtgcct | tccttgaccc | 3060 |
| tggaaggtgc | cactcccact | gtcctttcct | aataaaatga | ggaaattgca | tcgcattgtc | 3120 |
| tgagtaggtg | tcattctatt | ctggggggtg | gggtggggca | gcacagcaag | ggggaggatt | 3180 |
| gggaagacaa | tagcaggcat | gctggggatg | cggtgggctc | tatgggtacc | cagggccgca | 3240 |
| taacttcgta | taatgtatgc | tatacgaagt | tataagatct | gtactgaaat | gtgtgggcgt | 3300 |
| ggcttaaggg | tgggaaagaa | tatataaggt | gggggtctta | tgtagttttg | tatctqtttt | 3360 |
| gcagcagccg | ccgccgccat | gagcaccaac | tcgtttgatg | gaagcattqt | gageteatat | 3420 |
|            |            |            | gtgcgtcaga |            |            |      |
| gatggtcgcc | ccgtcctgcc | cgcaaactct | actaccttga | cctacgagac | cqtqtctqqa | 3540 |
| acgccgttgg | agactgcagc | ctccgccgcc | gcttcagccg | ctqcaqccac | caccacada  | 3600 |
| attgtgactg | actttgcttt | cctgagcccg | cttgcaagca | gtgcagcttc | ccqttcatcc | 3660 |
| gcccgcgatg | acaagttgac | ggctcttttg | gcacaattgg | attetttgac | ccqqqaactt | 3720 |
| aatgtcgttt | ctcagcagct | gttggatctg | cgccagcagg | tttctqccct | gaaggettee | 3780 |
| tcccctccca | atgcggttta | aaacataaat | aaaaaaccag | actctqtttq | gatttggatc | 3840 |
| aagcaagtgt | cttgctgtct | ttatttaggg | gttttgcgcg | cacaataaac | ccqqqaccaq | 3900 |
| cggtctcggt | cgttgagggt | cctgtgtatt | ttttccagga | cataataaaa | gtgactctgg | 3960 |
| atgttcagat | acatgggcat | aagcccgtct | ctggggtgga | qqtaqcacca | ctgcagaget | 4020 |
| tcatgctgcg | gggtggtgtt | gtagatgatc | cagtcgtagc | aggagcgctg | gacataatac | 4080 |
| ctaaaaatgt | ctttcagtag | caagctgatt | gccaggggca | ggcccttggt | gtaagtgttt | 4140 |
| acaaagcggt | taagctggga | tgggtgcata | cgtggggata | tgagatgcat | cttqqactqt | 4200 |
| atttttaggt | tggctatgtt | cccagccata | tccctccggg | gattcatgtt | gtgcagaacc | 4260 |
| accagcacag | tgtatccggt | gcacttggga | aatttgtcat | gtagettaga | aggaaatgcg | 4320 |
| tggaagaact | tggagacgcc | cttgtgacct | ccaagatttt | ccatqcattc | gtccataatg | 4380 |
| atggcaatgg | gcccacgggc | ggcggcctgg | gcgaagatat | ttctqqqatc | actaacqtca | 4440 |
| tagttgtgtt | ccaggatgag | atcgtcatag | gccattttta | caaaqcqcqq | acadagata  | 4500 |
| ccagactgcg | gtataatggt | tccatccggc | ccaggggcgt | agttaccctc | acagatttgc | 4560 |
| atttcccacg | ctttgagttc | agatggggg  | atcatgtcta | cctacaaaac | gatgaagaaa | 4620 |
| acggtttccg | gggtaggga  | gatcagctqq | gaagaaagca | ggttcctgag | cagctgcgac | 4680 |
| ttaccgcagc | cggtgggccc | gtaaatcaca | cctattaccg | gctgcaactg | gtagttaaga | 4740 |
|            |            |            | ggggccactt |            |            |      |
| cgcatgtttt | ccctgaccaa | atccccaga  | aggcgctcgc | cacccaacaa | taggagttet | 4860 |
|            |            |            | agaccgtccg |            |            |      |
|            |            |            | agctcggtca |            |            |      |
| tccagcatat | ctcctcattt | cacagattaa | ggcggctttc | gctgtacggc | agtagtcggt | 5040 |
| getegtecag | acqqqccaqq | gtcatgtctt | tccacgggcg | cagggtcche | atcaacataa | 5100 |
|            |            |            | gctgcgcgct |            |            |      |
|            |            |            | cttcgccctg |            |            |      |
| tgaccatgot | gtcatagtcc | agcccctccq | cggcgtggcc | cttggcacac | agettgeech | 5280 |
| tggaggaggc | gccgcacqaq | gagcagtaca | gacttttgag | ggcgtagagc | ttagacacaa | 5340 |
|            |            |            |            |            | 222-2-2-   |      |

```
gaaataccga ttccggggag taggcatccg cgccgcaggc cccgcagacg gtctcgcatt 5400
ccacgagcca ggtgagctct ggccgttcgg ggtcaaaaac caggtttccc ccatgctttt 5460
tgatgcgttt cttacctctg gtttccatga gccggtgtcc acgctcggtg acgaaaaggc 5520
tgtccgtgtc cccgtataca gacttgagag gcctgtcctc gagcggtgtt ccgcggtcct 5580
cetegtatag aaacteggae cactetgaga caaaggeteg egtecaggee ageacgaagg 5640
aggetaagtg ggaggggtag eggtegttgt ceactagggg gteeactege teeagggtgt 5700
gaagacacat gtcgccctct tcggcatcaa ggaaggtgat tqqtttqtaq qtqtaqqcca 5760
cgtgaccggg tgttcctgaa ggggggctat aaaagggggt gggggcgcqt tcqtcctcac 5820
tetetteege ategetgtet gegagggeea getgttgggg tgagtegaeg egaggetgga 5880
tggccttccc cattatgatt cttctcgctt ccggcggcat cgggatgccc gcgttgcagg 5940
ccatgctgtc caggcaggta gatgacgacc atcagggaca gcttcaaggc cagcaaaagg 6000
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccqc cccctqacq 6060
agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 6120
accaggogtt tececetgga ageteceteg tgegetetee tgtteegace etgeegetta 6180
coggatacct gtocgcottt ctcccttogg gaagogtggo gotttotoat agotcacgot 6240
gtaggtatet cagtteggtg taggtegtte getecaaget gggetgtgtg caeqaacee 6300
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa 6360
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 6420
taggeggtgc tacagagttc ttgaagtggt ggcctaacta eggctacact agaaggacag 6480
tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 6540
gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 6600
cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 6660
agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 6720
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 6780
cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 6840
ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct 6900
taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt 6960
tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat 7020
ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta 7080
atagtttgcg caacgttgtt gccattgctg caggcatcgt ggtgtcacgc tcgtcgtttg 7140
gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tcccccatgt 7200
tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccq 7260
cagtgttatc actcatggtt atggcagcac tgcataattc tcttactqtc atgccatccq 7320
taagatgett ttetgtgaet ggtgagtaet caaccaagte attetgagaa tagtgtatge 7380
ggcgaccgag ttgctcttgc ccggcgtcaa cacgggataa taccgcgcca catagcagaa 7440
ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac 7500
cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct tcagcatctt 7560
ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg 7620
gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa 7680
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt taqaaaaata 7740
aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca 7800
ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtcttcaag 7860
aattgtt
```

```
<210> 43
<211> 7870
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pAdApt containing HIV genes
<400> 43
```

|            |            | tgtttgacag |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            | atgagggggt |            |            |            |      |
|            |            | gtgtggcgga |            |            |            |      |
|            |            | aaagtgacgt |            |            |            |      |
| gacaattttc | gcgcggtttt | aggcggatgt | tgtagtaaat | ttgggcgtaa | ccgagtaaga | 300  |
| tttggccatt | ttcgcgggaa | aactgaataa | gaggaagtga | aatctgaata | attttgtgtt | 360  |
| actcatagcg | cgtaatattt | gtctagggcc | gcggggactt | tgaccgttta | cgtggagact | 420  |
|            |            | gtgttttccg |            |            |            |      |
|            |            | gcactggcct |            |            |            |      |
|            |            | tattggctca |            |            |            |      |
|            |            | gtaatcaatt |            |            |            |      |
|            |            | tacggtaaat |            |            |            |      |
|            |            | gacgtatgtt |            |            |            |      |
|            |            | tttacggtaa |            |            |            |      |
|            |            | tattgacgtc |            |            |            |      |
|            |            | ggactttcct |            |            |            |      |
|            |            | gttttggcag |            |            |            |      |
|            |            |            |            |            |            |      |
|            |            | ccacccatt  |            |            |            |      |
|            |            | atgtcgtaac |            |            |            |      |
|            |            | ctatataagc |            |            |            |      |
|            |            | ttttgacctc |            |            |            |      |
|            |            | gtgatcagat |            |            |            |      |
| agaagtacca | gcacctgtgg | cgctggggct | ggcgctgggg | caccatgctg | ctgggcatgc | 1380 |
|            |            | gagaagctgt |            |            |            |      |
| ggaaggaggc | caccaccacc | ctgctctgcg | ccagcgacgc | caaggcctac | gacaccgagg | 1500 |
|            |            | cacgcctgcg |            |            |            |      |
| tgctggtgaa | cgtgaccgag | aacttcgaca | tgtggaagaa | cgacatggtg | gagcagatgc | 1620 |
| acgaggacat | catcagcctg | tgggaccaga | gcctgaagcc | ctgcgtgaag | ctgacccccc | 1680 |
| tgtgcgtgag | cctgaagtgc | accgacctga | agaacgacac | caacaccaac | agcagcagcg | 1740 |
|            |            | ggcgagatca |            |            |            |      |
|            |            | gcctgcccca |            |            |            |      |
|            |            | gccatcctga |            |            |            |      |
|            |            | accgtgcagt |            |            |            |      |
|            |            | agcctggccg |            |            |            |      |
|            |            | atcatcgtgc |            |            |            |      |
|            |            | cgcaagagca |            |            |            |      |
|            |            | gacatccgcc |            |            |            |      |
|            |            | atcgtgatca |            |            |            |      |
|            |            | ggcggcgacc |            |            |            |      |
|            |            | aacagcgcta |            |            |            |      |
|            |            | aaggccatgt |            |            |            |      |
|            |            | ctgctgctga |            |            |            |      |
|            |            | ggcggcgaca |            |            |            |      |
|            |            | gageceetgg |            |            |            |      |
|            |            | agcggcatcg |            |            |            |      |
|            |            |            |            |            |            |      |
|            |            | ctgcagctga |            |            |            |      |
|            |            | tacctgaagg |            |            |            |      |
|            |            | gaccgcgaga |            |            |            |      |
|            |            | cagcaggaga |            |            |            |      |
|            |            | tggttcaaca |            |            |            |      |
|            |            | cagccatctg |            |            |            |      |
|            |            | actgtccttt |            |            |            |      |
|            |            | attctggggg |            |            |            |      |
|            |            | catgctgggg |            |            |            |      |
| gcataacttc | gtataatgta | tgctatacga | agttataaga | tctgtactga | aatgtgtggg | 3300 |

| _          |            |            |            |             |            |      |
|------------|------------|------------|------------|-------------|------------|------|
| cgtggcttaa | gggtgggaaa | gaatatataa | ggtgggggtc | ttatgtagtt  | ttgtatetgt | 3360 |
|            |            |            | aactcgtttg |             |            |      |
| tatttgacaa | cgcgcatgcc | cccatgggcc | ggggtgcgtc | agaatgtgat  | gggctccagc | 3480 |
| attgatggtc | gccccgtcct | gcccgcaaac | tctactacct | tgacctacga  | gaccgtgtct | 3540 |
| ggaacgccgt | tggagactgc | agcctccgcc | gccgcttcag | ccgctgcagc  | caccgcccgc | 3600 |
| gggattgtga | ctgactttgc | tttcctgagc | ccgcttgcaa | gcagtgcagc  | ttcccgttca | 3660 |
| teegeeegeg | atgacaagtt | gacggctctt | ttggcacaat | tggattcttt  | gacccgggaa | 3720 |
| cttaatgtcg | tttctcagca | gctgttggat | ctgcgccagc | aggtttctgc  | cctgaaggct | 3780 |
| tectecete  | ccaatgcggt | ttaaaacata | aataaaaac  | cagactctgt  | ttggatttgg | 3840 |
| atcaagcaag | tatcttacta | tctttattta | ggggttttgc | qcqcgcggta  | ggcccgggac | 3900 |
| cagcogtete | aatcattaaa | aatcctatat | atttttcca  | ggacgtggta  | aaggtgactc | 3960 |
| tagatatta  | gatacatggg | cataagcccg | tctctggggt | ggaggtagca  | ccactgcaga | 4020 |
|            |            |            | atccagtcgt |             |            |      |
| taataaaa   | tatatttaa  | tagcaagetg | attgccaggg | gcaggcctt   | gatataaata | 4140 |
| thtosoogs  | cattanasta | castagetes | atacgtgggg | atatgagatg  | catcttqqac | 4200 |
| Litataaage | ggttaagttg | ggacgggcgc | atacgcgggg | acacgagacg  | attataceae | 1260 |
| egeacettea | ggttggetat | geteetagee | atatecetee | ggggacccac  | agaagaaat  | 4320 |
| accaccagca | eagtgtatee | ggtgeaettg | ggaaatttgt | thteestees  | tteateasts | 4320 |
| gcgtggaaga | acttggagac | gecettgtga | cctccaagat | tittedatgea | atasatasaa | 4360 |
| atgatggcaa | tgggcccacg | aacaacaacc | tgggcgaaga | tatttetggg  | accaccaacg | 4440 |
| tcatagttgt | gttccaggat | gagatcgtca | taggccattt | ttacaaagcg  | cgggcggagg | 4500 |
| gtgccagact | gcggtataat | ggttccatcc | ggcccagggg | cgtagttacc  | ctcacagatt | 4560 |
| tgcatttccc | acgctttgag | ttcagatggg | gggatcatgt | ctacctgcgg  | ggcgatgaag | 4620 |
| aaaacggttt | ccggggtagg | ggagatcagc | tgggaagaaa | gcaggttcct  | gagcagctgc | 4680 |
| gacttaccgc | agccggtggg | cccgtaaatc | acacctatta | ccggctgcaa  | ctggtagtta | 4740 |
| agagagctgc | agctgccgtc | atccctgagc | aggggggcca | cttcgttaag  | catgtccctg | 4800 |
| actcgcatgt | tttccctgac | caaatccgcc | agaaggcgct | cgccgcccag  | cgatagcagt | 4860 |
| tcttgcaagg | aagcaaagtt | tttcaacggt | ttgagaccgt | ccgccgtagg  | catgcttttg | 4920 |
| agcgtttgac | caagcagttc | caggcggtcc | cacagctcgg | tcacctgctc  | tacggcatct | 4980 |
| cgatccagca | tatctcctcg | tttcgcgggt | tggggcggct | ttcgctgtac  | ggcagtagtc | 5040 |
| ggtgctcgtc | cagacgggcc | agggtcatgt | ctttccacgg | gcgcagggtc  | ctcgtcagcg | 5100 |
| tagtctgggt | cacggtgaag | gggtgcgctc | cgggctgcgc | gctggccagg  | gtgcgcttga | 5160 |
| ggctggtcct | gctggtgctg | aagcgctgcc | ggtcttcgcc | ctgcgcgtcg  | gccaggtagc | 5220 |
| atttgaccat | ggtgtcatag | tccagcccct | ccgcggcgtg | gcccttggcg  | cgcagcttgc | 5280 |
| ccttqqaqqa | qqcqccqcac | gaggggcagt | gcagactttt | gagggcgtag  | agcttgggcg | 5340 |
| cgagaaatac | cgattccggg | gagtaggcat | ccgcgccgca | ggccccgcag  | acggtctcgc | 5400 |
| attccacgag | ccaggtgagc | tctaaccatt | cggggtcaaa | aaccaggttt  | ccccatqct  | 5460 |
| ttttgatgcg | tttcttacct | ctggtttcca | tgagccggtg | tecaegeteg  | qtqacqaaaa | 5520 |
| gactatecat | gtccccgtat | acagacttga | gaggcctgtc | ctcgagcggt  | attccacaat | 5580 |
| cctcctcata | tagaaactcg | gaccactetg | agacaaaggc | tegegteeag  | gccagcacga | 5640 |
| aggaggetaa | atagaaaaaa | tagcggtcgt | tgtccactag | aggatecact  | cactccagga | 5700 |
| totoaagaca | catatacaca | tetteggeat | caaggaaggt | gattogtttg  | taggtgtagg | 5760 |
| ccacataaca | gaatatteet | daaddddac  | tataaaaggg | aataaaaaca  | cattcatcct | 5820 |
| cactetette | cacatcacta | tetacaaaaa | ccagctgttg | gaataaatca  | acacaaaact | 5880 |
| gastgacett | cccattata  | attettetea | cttccggcgg | catcoggato  | cccacattac | 5940 |
| aggeggeeee | atcasagesa | ateceteeg  | accatcaggg | acadetteaa  | ggcgagcaaa | 6000 |
| aggecatget | gcccaggcag | gragargarg | tggcgttttt | ccataggggg  | cacccccta  | 6060 |
| aggecaggaa | ccgcaaaaag | geegegeege | agaggeeee  | 222666262   | ggagtataa  | 6120 |
| acgageacea | caaaaaccga | egeteaagte | agaggtggcg | hashabbasa  | ggactataaa | 6100 |
| gataccagge | gtttcccct  | ggaageteee | tegtgegete | agagettest  | accordings | 6740 |
| ttaccggata | cctgtccgcc | CCCCCCCC   | cgggaagcgt | ggcgcttct   | catageteae | 6200 |
| gctgtaggta | tctcagttcg | gtgtaggtcg | ttcgctccaa | gergggergt  | gegeaegaae | 6300 |
|            |            |            | ccggtaacta |             |            |      |
| taagacacga | cttatcgcca | ctggcagcag | ccactggtaa | caggattage  | agagcgaggt | 6420 |
| atgtaggcgg | tgctacagag | ttettgaagt | ggtggcctaa | ctacggctac  | actagaagga | 6480 |
| cagtatttgg | tatctgcgct | ctgctgaagc | cagttacctt | cggaaaaaga  | gttggtagct | 6540 |
| cttgatccgg | caaacaaacc | accgctggta | gcggtggttt | ttttgtttgc  | aagcagcaga | 6600 |

```
ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg 6660
ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct 6720
tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 6780
aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctqtc 6840
tatttegtte atceatagtt geetgactee eegtegtgta gataactaeg ataegggagg 6900
gettaccate tggccccagt getgcaatga taccgcgaga cccacgetca ccqqctccaq 6960
atttatcago aataaaccag ccagooggaa gggoogagog cagaagtggt cotqcaactt 7020
tatecgeete catecagtet attaattgtt geegggaage tagaqtaaqt aqtteqeeaq 7080
ttaatagttt gegeaaegtt gttgeeattg etgeaggeat egtggtgtea egetegtegt 7140
ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca 7200
tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgq 7260
ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat 7320
ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta 7380
tgcggcgacc gagttgctct tgcccggcgt caacacggga taataccgcg ccacatagca 7440
gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct 7500
taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcaqcat 7560
cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat qccqcaaaaa 7620
agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt 7680
gaagcattta tcagggttat tgtctcatga gcggatacat atttqaatgt atttaqaaaa 7740'
ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa 7800
ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtcttc 7860
aagaattgtt
```

<210> 44 <211> 7885 <212> DNA <213> Artificial Sequence <220> <223> plasmid pAdApt containing HIV genes ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60 ggattgaagc caatatgata atgaggggt ggagtttgtg acgtggcgcg gggcgtggga 120 acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180 , tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240 gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300 tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttqtqtt 360 actcatageg egtaatattt gtetagggee geggggaett tgacegttta eqtqqaqact 420 cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480 tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540 tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600 ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660 gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720 cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780 tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840 catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900 geocagtaca tgacettatg ggaettteet acttggeagt acatetacgt attaqteate 960 gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020 tcacggggat ttccaagtct ccaccccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080 aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140 aggogtgtac ggtgggaggt ctatataagc agagotogtt tagtgaaccg tcagatogcc 1200 tggagacgec atccacgetg ttttgacete catagaagac accgggaccg atccageete 1260

|            |            |            |            |             | cgcgtgaagg |      |
|------------|------------|------------|------------|-------------|------------|------|
| agaagtacca |            |            |            |             |            |      |
|            |            |            |            |             | gtgcccgtgt |      |
|            |            |            |            |             | gacaccgagg |      |
|            |            |            |            |             | caggaggtgg |      |
|            |            |            |            |             | gagcagatgc |      |
|            |            |            |            |             | ctgacccccc |      |
| tgtgcgtgag | cctgaagtgc | accgacctga | agaacgacac | caacaccaac  | agcagcagcg | 1740 |
| gccgcatgat | catggagaag | ggcgagatca | agaactgcag | cttcaacatc  | agcaccagca | 1800 |
| tccgcggcaa | ggtgcagaag | gagtacgcct | tcttctacaa | gctggacatc  | atccccatcg | 1860 |
| acaacgacac | caccagctac | agcctgacca | gctgcaacac | cagcgtgatc  | acccaggcct | 1920 |
|            |            |            |            |             | ggcttcgcca |      |
|            |            |            |            |             | gtgagcaccg |      |
| tgcagtgcac | ccacggcatc | cgccccgtgg | tgagcaccca | gctgctgctg  | aacggtagcc | 2100 |
| tggccgagga | ggaggtggtg | atccgcagcg | ctaacttcgc | cgacaacgcc  | aaggtgatca | 2160 |
|            |            |            |            |             | cactgcaacc |      |
|            |            |            |            |             | cgcgagcagt |      |
|            |            |            |            |             | atcgtgaccc |      |
|            |            |            |            |             | ccctgccgca |      |
| tcaagcagat | catcaacatg | tggcagaagg | tgggcaaggc | catgtacgcc  | cccccatca  | 2460 |
|            |            |            |            |             | gacggcggca |      |
|            |            |            |            |             | gacaactggc |      |
| gcagcgagct | gtacaagtac | aaggtggtga | agatcgagcc | cctgggcgtg  | gccccacca  | 2640 |
|            |            |            |            |             | cagcagaaca |      |
| acctgctgcg | cgccatcgag | gcccagcagc | acctgctgca | gctgaccgtg  | tggggcatca | 2760 |
| agcagctgca | ggcccgcacc | ctggccgtgg | agcgctacct | gaaggaccag  | cagcageteg | 2820 |
|            |            |            |            |             | aactacacca |      |
|            |            |            |            |             | gagcaggagc |      |
|            |            |            |            |             | aactggctgt |      |
| ggtgaggatc | cagatctgct | gtgccttcta | gttgccagcc | atctgttgtt  | tgcccctccc | 3060 |
| ccgtgccttc | cttgaccctg | gaaggtgcca | ctcccactgt | cctttcctaa  | taaaatgagg | 3120 |
| aaattgcatc | gcattgtctg | agtaggtgtc | attctattct | aaaaaagtaaa | gtggggcagc | 3180 |
|            |            |            |            |             | gtgggctcta |      |
|            |            |            |            |             | taagatctgt |      |
|            |            |            |            |             | gggtcttatg |      |
|            |            |            |            |             | gtttgatgga |      |
|            |            |            |            |             | gcgtcagaat |      |
|            |            |            |            |             | taccttgacc |      |
| tacgagaccg | tgtctggaac | gccgttggag | actgcagcct | ccaccaccac  | ttcagccgct | 3600 |
| gcagccaccg | cccgcgggat | tgtgactgac | tttgctttcc | tgagcccgct  | tgcaagcagt | 3660 |
| gcagettece | gttcatccgc | ccgcgatgac | aagttgacgg | ctcttttggc  | acaattggat | 3720 |
| tctttgaccc | gggaacttaa | tgtcgtttct | cagcagctgt | tggatctgcg  | ccagcaggtt | 3780 |
|            |            |            |            |             | aaaaccagac |      |
| tctgtttgga | tttggatcaa | gcaagtgtct | tgctgtcttt | atttaggggt  | tttgcgcgcg | 3900 |
| cggtaggccc | gggaccagcg | gtctcggtcg | ctgagggtcc | tgtgtattt   | ttccaggacg | 3960 |
|            |            |            |            |             | ggggtggagg |      |
| tagcaccact | gcagagette | atgetgeggg | grggrgrrgr | agatgatcca  | gtcgtagcag | 4080 |
|            |            |            |            |             | caggggcagg |      |
|            |            |            |            |             | tggggatatg |      |
|            |            |            |            |             | cctccgggga |      |
|            |            |            |            |             | tttgtcatgt |      |
|            |            |            |            |             | aagattttcc |      |
|            |            |            |            |             | gaagatattt |      |
| ctgggatcac | caacgtcata | geegegeece | ayyargagar | cgccataggc  | catttttaca | 4500 |
| aagcgcgggc | ggagggtgcc | agactgcggt | acaacggttc | cateeggeee  | aggggcgtag | 4560 |

| ttaccctcac | agatttgcat | ttcccacgct | ttgagttcag | atggggggat | catgtctacc | 4620  |
|------------|------------|------------|------------|------------|------------|-------|
| tgcggggcga | tgaagaaaac | ggtttccggg | gtaggggaga | tcagctggga | agaaagcagg | 4680  |
| ttcctgagca | gctgcgactt | accgcagccg | gtgggcccgt | aaatcacacc | tattaccggc | 4740  |
| tgcaactggt | agttaagaga | gctgcagctg | ccgtcatccc | tgagcagggg | ggccacttcg | 4800  |
| ttaagcatgt | ccctgactcg | catgttttcc | ctgaccaaat | ccgccagaag | gcgctcgccg | 4860  |
| cccagcgata | gcagttcttg | caaggaagca | aagtttttca | acggtttgag | accgtccgcc | 4920  |
| gtaggcatgc | ttttgagcgt | ttgaccaagc | agttccaggc | ggtcccacag | ctcggtcacc | 4980  |
| tgctctacgg | catctcgatc | cagcatatct | cctcgtttcg | cgggttgggg | cggctttcgc | 5040  |
| tgtacggcag | tagtcggtgc | tcgtccagac | gggccagggt | catgtctttc | cacgggcgca | 5100  |
| gggtcctcgt | cagcgtagtc | tgggtcacgg | tgaaggggtg | cgctccgggc | tgcgcgctgg | 5160  |
| ccagggtgcg | cttgaggctg | gtcctgctgg | tgctgaagcg | ctgccggtct | tegeeetgeg | 5220  |
| cgtcggccag | gtagcatttg | accatggtgt | catagtccag | cccctccgcg | gcgtggccct | 5280  |
| tggcgcgcag | cttgcccttg | gaggaggcgc | cgcacgaggg | gcagtgcaga | cttttgaggg | 5340  |
| cgtagagctt | gggcgcgaga | aataccgatt | ccggggagta | ggcatccgcg | ccgcaggccc | 5400  |
| cgcagacggt | ctcgcattcc | acgagccagg | tgagetetgg | ccattcagaa | tcaaaaacca | 5460  |
| ggtttcccc  | atgctttttg | atgcgtttct | tacctctggt | ttccatgage | cggtgtccac | 5520  |
| gctcggtgac | gaaaaggctg | tecgtgtece | cqtatacaqa | cttgagaggc | ctgtcctcga | 5580  |
| gcggtgttcc | gcggtcctcc | tcgtatagaa | actoggacca | ctctgagaca | aaggctcgcg | 564.0 |
| tccaggccag | cacqaaqqaq | qctaaqtqqq | aggggtagcg | atcattatca | actagggggt | 5700  |
| ccactcqctc | cagggtgtga | agacacatgt | caccetette | ggcatcaagg | aaggtgattg | 5760  |
| atttatagat | gtaggccacq | tgaccgggtg | ttcctgaagg | ggggctataa | aagggggtgg | 5820  |
| gagcacattc | gtcctcactc | tcttccccat | cactatetae | gagggccagc | tgttggggtg | 5880  |
| agtcgacgcg | aggctggatg | gccttcccca | ttatgattct | tetegettee | ggcggcatcg | 5040  |
| ggatgcccgc | attacagacc | atactatcca | gacagataga | tgacgaccat | cagggacagc | 5010  |
| ttcaaggcca | gcaaaaggcc | aggaaccgta | aaaaaaccac | attactaca  | tttttccata | 6060  |
| gachecaece | ccctgacgag | catcacaaaa | atcascactc | antragag   | tggcgaaacc | 6120  |
| caacagact  | ataaagatac | caggggtttc | ccctagaaa  | ctocateata | cgctctcctg | 6100  |
| ttccgaccct | accacttace | quatacetet | ccacctttct | cccttccgcg | agegtggege | 6100  |
| tttctcatac | ctcacctat  | aggtatetea | atteastata | ccccccggga | tccaagctgg | 6240  |
| actatatace | caaaccccc  | atteaceea  | accostosos | ggtegteege | recaagetgg | 6300  |
| ttgagtgca  | cccateea   | geccagetes | accecteege | cccaccegge | aactatcgtc | 6360  |
| ttaggeceau | gaggtatga  | gacgataata | gecaetgge  | ageageeaee | ggtaacagga | 6420  |
| actacactac | aaggcacgca | thtachatat | cagageteet | gaagtggtgg | cctaactacg | 6480  |
| assasattas | taggacagca | teeggeatee | gegetetget | gaagecagee | accttcggaa | 6540  |
| tttggaage  | aceastrea  | cccggcaaac | aaaccaccgc | tggtageggt | ggttttttg  | 6600  |
| ctacaaacta | tancactana | taanaaaa   | aaggaccca  | agaagateet | ttgatctttt | 6660  |
| tataaaaaa  | cgacgereag | togaacgaaa | acceaegtta | agggatttg  | gtcatgagat | 6720  |
| accaaaaay  | taratarash | tagateette | taaattaaaa | acgaagetet | aaatcaatct | 6780  |
| tatasaaaat | cyaytaaact | tggtctgaca | gttaccaatg | cttaatcagt | gaggcaccta | 6840  |
| chagaga    | agaggaabta | egttcatcca | tagttgcctg | actccccgtc | gtgtagataa | 6900  |
| gatasagga  | taasasttta | beauciggee | ccagtgetge | aatgataccg | cgagacccac | 6960  |
| geceaeegge | cecayactta | ccagcaataa | accagecage | cggaagggcc | gagcgcagaa | 7020  |
| gragererge | aactttatee | gectecatee | agretattaa | ttgttgccgg | gaagctagag | 7080  |
| taagtagtte | gecagetaat | agtttgcgca | acgttgttgc | cattgctgca | ggcatcgtgg | 7140  |
| tgtcacgctc | gregrergge | atggetteat | tcagctccgg | ttcccaacga | tcaaggcgag | 7200  |
| ttacatgate | cccatgttg  | cgcaaaaaag | cggttagctc | cttcggtcct | ccgatcgttg | 7260  |
| tcagaagtaa | grrggccgca | gtgttatcac | tcatggttat | ggcagcactg | cataattctc | 7320  |
| ttactgtcat | gccatccgta | agatgctttt | ctgtgactgg | tgagtactca | accaagtcat | 7380  |
| tctgagaata | gtgtatgcgg | cgaccgagtt | gctcttgccc | ggcgtcaaca | cgggataata | 7440  |
| ccgcgccaca | cagcagaact | ttaaaagtgc | tcatcattgg | aaaacgttct | tcggggcgaa | 7500  |
| aactctcaag | gatettaceg | ccgctgagat | ccagttcgat | gtaacccact | cgtgcaccca | 7560  |
| actgatcttc | agcatctttt | actttcacca | gcgtttctgg | gtgagcaaaa | acaggaaggc | 7620  |
| aaaatgccgc | aaaaaaggga | ataagggcga | cacggaaatg | ttgaatactc | atactcttcc | 7680  |
| tttttcaata | ttattgaagc | atttatcagg | gttattgtct | catgagcgga | tacatatttg | 7740  |
| aatgtattta | gaaaaataaa | caaatagggg | ttccgcgcac | atttccccga | aaagtgccac | 7800  |
| ctgacgtcta | agaaaccatt | attatcatga | cattaaccta | taaaaatagg | cgtatcacga | 7860  |

ggccctttcg tcttcaagaa ttgtt

7885

```
<210> 45
<211> 7789
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pAdApt containing HIV genes
<400> 45
ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggt ggagtttgtg acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagegae ggatgtggea aaagtgaegt ttttggtgtg cgccqqtqta cacaqqaaqt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
actcatagcg cgtaatattt gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480
tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660
gagtteegeg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaaqtqtat 840
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 960
gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020
tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140
aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200
tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260
cgtcaccgtc gtcgacacgt gtgatcagat atcgcggccg ctctagaatg cgcgtgaagg 1320
agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380
tgatgatetg cagegecace gagaagetgt gggtgacegt gtactaegge gtgcccgtgt 1440
ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500
tgcacaacgt gtgggccacc cacgcctgcg tgcccaccga ccccaacccc caggaggtgg 1560
tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcgtgaag ctgaccccc 1680
tgtgcgtgag cctgaagtgc accgacgcta gcaccagctg caacaccagc gtgatcaccc 1740
aggeotgece caaggtgage ttegageeca tececaacea etactgegee ecegeegget 1800
tegecatect gaagtgeaag gacaagaagt teaaeggeaa gggeeeetge accaaegtga 1860
gcaccgtgca gtgcacccac ggcatccgcc ccgtggtgag cacccagctg ctgctgaacg 1920
gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcgccgac aacgccaagg 1980
tgatcategt geagetgaae gagagegtgg agateaactg caceegeget agegeecaet 2040
gcaacctgag ccgcgccaag tggaacgaca ccctgaacaa gatcgtgatc aagctgcgcg 2100
agcagttcgg caacaagacc atcgtgttca agcacagcag cggcggcgac cccgagatcg 2160
tgacccacag cttcaactgc ggcggcgagt tcttctactg caacagcacc cagctgttca 2220
acagcacetg gttcaacage acetggagea eegagggeag caacaacace gagggeageg 2280
acaccatcac cetgecetge egeateaage agateateaa eatgtggeag aaggtgggea 2340
aggccatgta cgccccccc atcagcggcc agatccgctg cagcagcaac atcaccggcc 2400
tgctgctgac ccgcgacggc ggcaacagca acaacgagag cgagatcttc cgtccgggcg 2460
gcggcgacat gcgcgacaac tggcgcagcg agctgtacaa gtacaaggtg gtgaagatcg 2520
```

|            |            |            |            |            | cagctgctga |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | cagcacctgc |      |
| tgcagctgac | cgtgtggggc | atcaagcagc | tgcaggcccg | caccctggcc | gtggagcgct | 2700 |
|            |            |            |            |            | atggagtggg |      |
|            |            |            |            |            | agccagaacc |      |
|            |            |            |            |            | ctgtggaact |      |
|            |            |            |            |            | tctagttgcc |      |
|            |            |            |            |            | gccactccca |      |
|            |            |            |            |            | tgtcattcta |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | aatagcaggc |      |
|            |            |            |            |            | tataatgtat |      |
|            |            |            |            |            | ggtgggaaag |      |
|            |            |            |            |            | cgccgccgcc |      |
|            |            |            |            |            | gcgcatgccc |      |
|            |            |            |            |            | ccccgtcctg |      |
| cccgcaaact | ctactacctt | gacctacgag | accgtgtctg | gaacgccgtt | ggagactgca | 3480 |
| gecteegeeg | ccgcttcagc | cgctgcagcc | accgcccgcg | ggattgtgac | tgactttgct | 3540 |
| ttcctgagcc | cgcttgcaag | tagtgcagct | tcccgttcat | ccgcccgcga | tgacaagttg | 3600 |
| acggctcttt | tggcacaatt | ggattctttg | acccgggaac | ttaatgtcgt | ttctcagcag | 3660 |
| ctgttggatc | tgcgccagca | ggtttctgcc | ctgaaggctt | cctccctcc  | caatgcggtt | 3720 |
|            |            |            |            |            | gtcttgctgt |      |
|            |            |            |            |            | gtcgttgagg |      |
|            |            |            |            |            | atacatgggc |      |
|            |            |            |            |            | cggggtggtg |      |
|            |            |            |            |            | gtctttcagt |      |
|            |            |            |            |            | gttaagctgg |      |
|            |            |            |            |            | gttggctatg |      |
|            |            |            |            |            | agtgtatccg |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | cttggagacg |      |
|            |            |            |            |            | gggcccacgg |      |
|            |            |            |            |            | ttccaggatg |      |
|            |            |            |            |            | cggtataatg |      |
|            |            |            |            |            | cgctttgagt |      |
|            |            |            |            |            | cggggtaggg |      |
|            |            |            |            |            | accaaraaac |      |
|            |            |            |            |            | gctgccgtca |      |
|            |            |            |            |            | ttccctgacc |      |
| aaatccgcca | gaaggcgctc | gccgcccagc | gatagcagtt | cttgcaagga | agcaaagttt | 4800 |
| ttcaacggtt | tgagaccgtc | cgccgtaggc | atgcttttga | gcgtttgacc | aagcagttcc | 4860 |
| aggcggtccc | acagctcggt | cacctgctct | acggcatctc | gatccagcat | atctcctcgt | 4920 |
|            |            |            |            |            | agacgggcca |      |
| gggtcatgtc | tttccacggg | cgcagggtcc | tcgtcagcgt | agtctgggtc | acggtgaagg | 5040 |
| ggtgcgctcc | gggctgcgcg | ctggccaggg | tgcgcttgag | gctggtcctg | ctggtgctga | 5100 |
| agcgctgccg | gtcttcgccc | tgcgcgtcgg | ccaggtagca | tttgaccatg | gtgtcatagt | 5160 |
|            |            |            |            |            | gcgccgcacg |      |
|            |            |            |            |            | gattccgggg |      |
|            |            |            |            |            | caggtgagct |      |
|            |            |            |            |            | ttcttacctc |      |
|            |            | _          |            |            | teccegtáta |      |
|            |            |            |            |            | agaaactcgg |      |
|            |            |            |            |            | tgggaggggt |      |
|            |            |            |            |            | atgtcgccct |      |
|            |            |            |            |            | ggtgttcctg |      |
|            |            |            |            |            | gcatcgctgt |      |
|            |            |            |            |            |            |      |
| ccycyagggc | cagoogtogg | ggtgagttga | cycyaggetg | gatggcccc  | cccattatga | 2020 |

```
ttettetege tteeggegge ategggatge eegegttgea ggecatgetg teeaqqeaqq 5880
tagatgacga ccatcaggga cagetteaag gecaqeaaaa qqecaqqaac eqtaaaaaqq 5940
ccgcgttgct ggcgtttttc cataggctcc gccccctqa cqaqcatcac aaaaatcqac 6000
gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg 6060
gaageteest egtgegetet cetgtteega ceetgeeget taceggatae etgteegeet 6120
ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg 6180
tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 6240
gegeettate eggtaactat egtettgagt ecaaceeggt aagacaegae ttategeeac 6300
tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt 6360
tettgaagtg gtggeetaae taeggetaea etagaaggae agtatttggt atetgegete 6420
tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 6480
ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 6540
ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac 6600
gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 6660
aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct qacaqttacc 6720
aatgettaat cagtgaggea cetateteag egatetgtet atttegttea tecatagttg 6780
cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg 6840
ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc 6900
cageeggaag ggeegagege agaagtggte etgeaacttt ateegeetee ateeaqteta 6960
ttaatbgttg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg 7020
ttgccattgc tgcaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct 7080
ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta 7140
getecttegg teeteegate gttgteagaa gtaagttgge egeagtgtta teacteatgg 7200
ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga 7260
ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg aqttqctctt 7320
gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 7380
ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccaqtt 7440
cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt 7500
ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga 7560
aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat cagggttatt 7620
gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 7680
gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa 7740
cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattgtt
                                                                  7789
```

```
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pAdApt containing HIV genes

<400> 46
ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttattt 60
ggattgaagc caatatgata atgaggggt ggagtttgtg acgtggcgc gggcgtggga 120
acggggcggg tgacqtagta gtgtggcga agtgtgatgt tgcaagtgt gcggaacaca 180
tgtaagcgac ggattggca aaagtgacgt ttttggtgt cgccggtgta cacaggaagt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgt 360
actcatagcg cgtaatattt gtctagggcc gcggggactt tgaccgtta cgtggagact 420
cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480
tatagtcagt acgtaccagt gcactggct tgtccaacat taccgccatg ttgacattga 600
```

<210> 46
<211> 7792

| A-A 1-1 A- |            |            |            | L          |            | cc0  |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            | tagttcatag |            |      |
|            |            |            |            | gctgaccgcc |            |      |
|            |            |            |            | cgccaatagg |            |      |
|            |            |            |            | tggcagtaca |            |      |
|            |            |            |            | aatggcccgc |            |      |
| gcccagtaca | tgaccttatg | ggactttcct | acttggcagt | acatctacgt | attagtcatc | 960  |
| gctattacca | tggtgatgcg | gttttggcag | tacatcaatg | ggcgtggata | gcggtttgac | 1020 |
| tcacggggat | ttccaagtct | ccaccccatt | gacgtcaatg | ggagtttgtt | ttggcaccaa | 1080 |
|            |            |            |            | cattgacgca |            |      |
|            |            |            |            | tagtgaaccg |            |      |
|            |            |            |            | accgggaccg |            |      |
|            |            |            |            | ctctagaatg |            |      |
|            |            |            |            | caccatgctg |            |      |
|            |            |            |            | gtactacggc |            |      |
|            |            |            |            | caaggcctac |            |      |
|            |            |            |            | ccccaaccc  |            |      |
|            |            |            |            |            |            |      |
|            |            |            |            | cgacatggtg |            |      |
|            |            |            |            | ctgcgtgaag |            |      |
|            |            |            |            | caacaccagc |            |      |
|            |            |            |            | ctactgcgcc |            |      |
| tegecatect | gaagtgcaag | gacaagaagt | tcaacggcaa | gggcccctgc | accaacgtga | 1860 |
| gcaccgtgca | gtgcacccac | ggcatccgcc | ccgtggtgag | cacccagctg | ctgctgaacg | 1920 |
|            |            |            |            | cttcgccgac |            |      |
|            |            |            |            | cacccgcccc |            |      |
|            |            |            |            | caccaccggc |            |      |
| gcgacatccg | ccaggcccac | tgcaacctga | gccgcgccaa | gtggaacgac | accctgaaca | 2160 |
| agatcgtgat | caagctgcgc | gagcagttcg | gcaacaagac | catcgtgttc | aagcacagca | 2220 |
| gcggcggcga | ccccgagatc | gtgacccaca | gcttcaactg | cggcggcgag | ttcttctact | 2280 |
|            |            |            |            | caacatgtgg |            |      |
|            |            |            |            | ctgcagcagc |            |      |
|            |            |            |            | gagcgagatc |            |      |
| acaacaacaa | catacacaac | aactggcgca | gegagetgta | caagtacaag | gtggtgaaga | 2520 |
| tcgagcccct | agacataace | cccaccaaqq | ccaagettae | cgtccaggcc | cqccaqctqc | 2580 |
|            |            |            |            | catcgaggcc |            |      |
|            |            |            |            | ccgcaccctg |            |      |
|            |            |            |            | ccacaccacc |            |      |
|            |            |            |            | cctgatcgag |            |      |
|            |            |            |            | caagtgggcc |            |      |
| accagcagga | gaagaacgag | taagtataat | gagagaga   | atctgctgtg | ccttctagtt | 2940 |
| actggcccaa | tattattaa  | cagecacaa  | tacetteett | gaccetagaa | cataccaete | 3000 |
| geeageeate | thattat    | actececceg | ttaattaa   | gaccctggaa | aggregates | 3060 |
| ccactgtcct | ttectaataa | aatgaggaaa | regeategea | ttgtctgagt | aggigicaci | 3120 |
|            |            |            |            | ggattgggaa |            |      |
|            |            |            |            | ccgcataact |            |      |
| tatgctatac | gaagttataa | gatetgtact | gaaatgtgtg | ggcgtggctt | aagggtggga | 3240 |
|            |            |            |            | gttttgcagc |            |      |
|            |            |            |            | catatttgac |            |      |
|            |            |            |            | gcattgatgg |            |      |
|            |            |            |            | ctggaacgcc |            |      |
| gcagcctccg | ccgccgcttc | agccgctgca | gccaccgccc | gcgggattgt | gactgacttt | 3540 |
| gctttcctga | gcccgcttgc | aagcagtgca | gcttcccgtt | catecgeeeg | cgatgacaag | 3600 |
| ttgacggctc | ttttggcaca | attggattct | ttgacccggg | aacttaatgt | cgtttctcag | 3660 |
| cagctgttgg | atctgcgcca | gcaggtttct | gccctgaagg | cttcctcccc | tcccaatgcg | 3720 |
| gtttaaaaca | taaataaaaa | accagactct | gtttggattt | ggatcaagca | agtgtcttgc | 3780 |
|            |            |            |            | accagcggtc |            |      |
|            |            |            |            | tctggatgtt |            |      |
|            |            |            |            |            | _          |      |

|   |            |            |            |            |            | ctgcggggtg |      |
|---|------------|------------|------------|------------|------------|------------|------|
|   |            |            |            |            |            | aatgtctttc |      |
|   |            |            |            |            |            | gcggttaagc |      |
|   |            |            |            |            |            | taggttggct |      |
|   | atgttcccag | ccatatccct | ccggggattc | atgttgtgca | gaaccaccag | cacagtgtat | 4200 |
|   | ccggtgcact | tgggaaattt | gtcatgtagc | ttagaaggaa | atgcgtggaa | gaacttggag | 4260 |
|   | acgcccttgt | gacctccaag | attttccatg | cattcgtcca | taatgatggc | aatgggccca | 4320 |
|   |            |            |            |            |            | gtgttccagg |      |
|   | atgagatcgt | cataggccat | ttttacaaag | cgcgggcgga | gggtgccaga | ctgcggtata | 4440 |
|   | atggttccat | ccggcccagg | ggcgtagtta | ccctcacaga | tttgcatttc | ccacgctttg | 4500 |
|   | agttcagatg | gggggatcat | gtctacctgc | ggggcgatga | agaaaacggt | ttccggggta | 4560 |
|   |            |            |            |            |            | gcagccggtg |      |
| • |            |            |            |            |            | gcagctgccg |      |
|   | tcatccctga | gcaggggggc | cacttcgtta | agcatgtccc | tgactcgcat | gttttccctg | 4740 |
|   |            |            |            |            |            | ggaagcaaag |      |
|   | tttttcaacg | gtttgagacc | gtccgccgta | ggcatgcttt | tgagcgtttg | accaagcagt | 4860 |
|   |            |            |            |            |            | catatctcct |      |
|   | cgtttcgcgg | gttggggcgg | ctttcgctgt | acggcagtag | teggtgeteg | tccagacggg | 4980 |
|   | ccagggtcat | gtctttccac | gggcgcaggg | tcctcgtcag | cgtagtctgg | gtcacggtga | 5040 |
|   | aggggtgcgc | tccgggctgc | gcgctggcca | gggtgcgctt | gaggctggtc | ctgctggtgc | 5100 |
|   | tgaagcgctg | ccggtcttcg | ccctgcgcgt | cggccaggta | gcatttgacc | atggtgtcat | 5160 |
|   | agtccagccc | ctccgcggcg | tggcccttgg | cgcgcagctt | gcccttggag | gaggcgccgc | 5220 |
|   |            |            |            |            |            | accgattccg |      |
|   | gggagtaggc | atccgcgccg | caggccccgc | agacggtctc | gcattccacg | agccaggtga | 5340 |
|   | gctctggccg | ttcggggtca | aaaaccaggt | ttcccccatg | ctttttgatg | cgtttcttac | 5400 |
|   |            |            |            |            |            | gtgtccccgt |      |
|   | atacagactt | gagaggcctg | tcctcgagcg | gtgttccgcg | gtcctcctcg | tatagaaact | 5520 |
|   | cggaccactc | tgagacaaag | gctcgcgtcc | aggccagcac | gaaggaggct | aagtgggagg | 5580 |
|   | ggtagcggtc | gttgtccact | agggggtcca | ctcgctccag | ggtgtgaaga | cacatgtcgc | 5640 |
|   |            |            |            |            |            | ccgggtgttc |      |
|   | ctgaaggggg | gctataaaag | aaaaraaaaa | cgcgttcgtc | ctcactctct | tccgcatcgc | 5760 |
|   | tgtctgcgag | ggccagctgt | tggggtgagt | cgacgcgagg | ctggatggcc | ttccccatta | 5820 |
|   | tgattcttct | cgcttccggc | ggcatcggga | tgcccgcgtt | gcaggccatg | ctgtccaggc | 5880 |
|   | aggtagatga | cgaccatcag | ggacagcttc | aaggccagca | aaaggccagg | aaccgtaaaa | 5940 |
|   | aggccgcgtt | gctggcgttt | ttccataggc | tecgccccc  | tgacgagcat | cacaaaaatc | 6000 |
| • | gacgctcaag | tcagaggtgg | cgaaacccga | caggactata | aagataccag | gcgtttcccc | 6060 |
|   | ctggaagctc | cctcgtgcgc | tctcctgttc | cgaccctgcc | gcttaccgga | tacctgtccg | 6120 |
|   | cctttctccc | ttcgggaagc | gtggcgcttt | ctcatagctc | acgctgtagg | tatctcagtt | 6180 |
|   | cggtgtaggt | cgttcgctcc | aagctgggct | gtgtgcacga | accccccgtt | cagcccgacc | 6240 |
|   |            |            |            |            |            | gacttatcgc |      |
|   |            |            |            |            |            | ggtgctacag |      |
|   |            |            |            |            |            | ggtatctgcg |      |
|   | ctctgctgaa | gccagttacc | ttcggaaaaa | gagttggtag | ctcttgatcc | ggcaaacaaa | 6480 |
|   |            |            |            |            |            | agaaaaaaag |      |
|   | gatctcaaga | agatcctttg | atcttttcta | cggggtctga | cgctcagtgg | aacgaaaact | 6600 |
|   |            |            |            |            |            | atccttttaa |      |
|   | attaaaaatg | aagttttaaa | tcaatctaaa | gtatatatga | gtaaacttgg | tctgacagtt | 6720 |
|   | _          |            |            |            |            | tcatccatag |      |
|   |            |            |            |            |            | tctggcccca |      |
|   | _          | -          |            |            |            | gcaataaacc |      |
|   |            |            |            |            |            | tccatccagt |      |
|   |            |            |            |            |            | trgcgcaacg |      |
|   |            |            |            |            |            | gcttcattca |      |
|   |            |            |            |            |            | aaaaaagcgg |      |
|   | ttagctcctt | cggtcctccg | atcgttgtca | gaagtaagtt | ggccgcagtg | ttatcactca | 7200 |
|   |            |            |            |            |            |            |      |

```
tggttatgge agcactgcat aattetetta etgteatge atcegtaaga tgettteteg 7260 tgactggtga gtacteaace aagteattet gagaatagtg tatgeggega eegagttget 7320 ettgeeegge gteaacaegg gataataceg egceacatag cagaacttta aaagtgetea 7380 teattggaaa acgttetteg gggegaaaac teteaaggat ettacegetg ttgagateea 7440 gttegatgta acceaetegt geaceeaace gatetteage atetttaet tteaecageg 7500 tttetgggtg agcaaaaca ggaaggeaaa atgeegeaaa aaagggaata agggegaaca 7560 ggaaatgttg aatacteata etetteett tteaatatta ttgaagcatt tateagggtt 7620 attgteeat teeecagaaa gtgeeacetg gatettaaga aaataaacaa ataggggtte 7680 egegeacatt teeecagaaa gtgeeacetg acgtetaaga aaceattatt ateatgacat 7740 taacetataa aaataggegt ateaegagge eetttegtet teaagaattg tt 7792
```

<210> 47
<211> 7828
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pAdApt containing HIV genes

<400> 47 ttaattaacc gcaattetea tgtttgacag ettateatea tcaataatat acettatttt 60 ggattgaagc caatatgata atgagggggt ggagtttgtg acgtggcgcg gggcgtggga 120 acqqqcqqq tgacqtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180 tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240 gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttggggcgtaa ccgagtaaga 300 tttqqccatt ttcqcqqqaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360 actcatagcg cgtaatattt gtctagggcc gcggggactt tgaccgttta cgtggagact 420 cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480 tataqtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540 tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600 ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660 gagtteegeg ttacataact taeggtaaat ggeeegeetg getgaeegee eaacgaeece 720 egeceattga egteaataat gaegtatgtt eecatagtaa egecaatagg gaettteeat 780 tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840 catatgccaa gtacgcccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900 gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 960 gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020 tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080 aatcaacqqq actttccaaa atqtcqtaac aactccgccc cattgacgca aatgggcggt 1140 aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200 tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260 cgtcaccgtc gtcgacacgt gtgatcagat atcgcggccg ctctagaatg cgcgtgaagg 1320 agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380 tgatgatetg cagegecace gagaagetgt gggtgacegt gtactacgge gtgcccgtgt 1440 ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500 tgcacaacgt gtgggccacc cacgcctgcg tgcccaccga ccccaacccc caggaggtgg 1560 tqctggtgaa cqtqaccqaq aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620 acqaqqacat catcaqcctq tqqqaccaga gcctgaagcc ctgcgtgaag ctgaccccc 1680 tqtgcgtgag cctgaaqtqc accgacgcta gcaagaactg cagcttcaac atcagcacca 1740 gcatccgcgg caaggtgcag aaggagtacg cettetteta caagetggae atcatcccca 1800 tegacaacga caccaccage tacageetga ecagetgeaa caccagegtg atcacccagg 1860 cctgccccaa ggtgagcttc gagcccatcc ccaaccacta ctgcgccccc gccggcttcg 1920 ccatcctgaa gtgcaaggac aagaagttca acggcaaggg cccctgcacc aacgtgagca 1980

| ccgtgcagtg | cacccacggc | atccgccccg  | tggtgagcac | ccagctgctg | ctgaacggta | 2040 |
|------------|------------|-------------|------------|------------|------------|------|
| gcctggccga | ggaggaggtg | gtgatccgca  | gcgctaactt | cgccgacaac | gccaaggtga | 2100 |
| tcatcgtgca | gctgaacgag | agcgtggaga  | tcaactgcac | ccgcgctagc | gcccactgca | 2160 |
| acctgagccg | cgccaagtgg | aacgacaccc  | tgaacaagat | cgtgatcaag | ctgcgcgagc | 2220 |
| agttcggcaa | caagaccatc | gtgttcaagc  | acagcagcgg | cggcgacccc | gagatcgtga | 2280 |
| cccacagett | caactgcggc | ggcgagttct  | tctactgcaa | cagcgctagc | ctgccctgcc | 2340 |
| gcatcaagca | gatcatcaac | atgtggcaga  | aggtgggcaa | ggccatgtac | gccccccca  | 2400 |
| tcagcggcca | gatccgctgc | agcaġcaaca  | tcaccggcct | gctgctgacc | cgcgacggcg | 2460 |
| gcaacagcaa | caacgagagc | gagatettee  | gtccgggcgg | cggcgacatg | cgcgacaact | 2520 |
| ggcgcagcga | gctgtacaag | tacaaggtgg  | tgaagatcga | gcccctgggc | gtggccccca | 2580 |
| ccaaggccaa | gcttaccgtc | caggcccgcc  | agctgctgag | cggcatcgtg | cagcagcaga | 2640 |
| acaacctgct | gcgcgccatc | gaggcccagc  | agcacctgct | gcagctgacc | gtgtggggca | 2700 |
| tcaagcagct | gcaggcccgc | accctggccg  | tggagcgcta | cctgaaggac | cagcagcagc | 2760 |
| tcgagcagat | ctggaaccac | accacctgga  | tggagtggga | ccgcgagatc | aacaactaca | 2820 |
| ccagcctgat | ccacagcctg | atcgaggaga  | gccagaacca | gcaggagaag | aacgagcagg | 2880 |
| agctgctgga | gctggacaag | tgggccagcc  | tgtggaactg | gttcaacatc | accaactggc | 2940 |
| tgtggtgagg | atccagatct | gctgtgcctt  | ctagttgcca | gccatctgtt | gtttgcccct | 3000 |
| cccccgtgcc | ttccttgacc | ctggaaggtg  | ccactcccac | tgtcctttcc | taataaaatg | 3060 |
| aggaaattgc | atcgcattgt | ctgagtaggt  | gtcattctat | tctggggggt | ggggtggggc | 3120 |
| agcacagcaa | gggggaggat | tgggaagaca  | atagcaggca | tgctggggat | gcggtgggct | 3180 |
| ctatgggtac | ccagggccgc | ataacttcgt  | ataatgtatg | ctatacgaag | ttataagatc | 3240 |
| tgtactgaaa | tgtgtgggcg | tggcttaagg. | gtgggaaaga | atatataagg | tgggggtctt | 3300 |
| atgtagtttt | gtatctgttt | tġcagcagcc  | gccgccgcca | tgagcaccaa | ctcgtttgat | 3360 |
| ggaagcattg | tgagctcata | tttgacaacg  | cgcatgcccc | catgggccgg | ggtgcgtcag | 3420 |
| aatgtgatgg | gctccagcat | tgatggtcgc  | cccgtcctgc | ccgcaaactc | tactaccttg | 3480 |
| acctacgaga | ccgtgtctgg | aacgccgttg  | gagactgcag | cctccgccgc | cgcttcagcc | 3540 |
| gctgcagcca | ccgcccgcgg | gattgtgact  | gactttgctt | tcctgagccc | gcttgcaagc | 3600 |
| agtgcagctt | cccgttcatc | cgcccgcgat  | gacaagttga | cggctctttt | ggcacaattg | 3660 |
| gattctttga | cccgggaact | taatgtcgtt  | tctcagcagc | tgttggatct | gcgccagcag | 3720 |
| gtttctgccc | tgaaggcttc | ctcccctccc  | aatgcggttt | aaaacataaa | taaaaaacca | 3780 |
| gactctgttt | ggatttggat | caagcaagtg  | tcttgctgtc | tttatttagg | ggttttgcgc | 3840 |
|            |            | gcggtctcgg  |            |            |            |      |
| acgtggtaaa | ggtgactctg | gatgttcaga  | tacatgggca | taagcccgtc | tctggggtgg | 3960 |
|            |            | ttcatgctgc  |            |            |            |      |
| caggagcgct | gggcgtggtg | cctaaaaatg  | tctttcagta | gcaagctgat | tgccaggggc | 4080 |
| aggcccttgg | tgtaagtgtt | tacaaagcgg  | ttaagctggg | atgggtgcat | acgtggggat | 4140 |
| atgagatgca | tettggaetg | tatttttagg  | ttggctatgt | tcccagccat | atccctccgg | 4200 |
| ggattcatgt | tgtgcagaac | caccagcaca  | gtgtatccgg | tgcacttggg | aaatttgtca | 4260 |
| tgtagcttag | aaggaaatgc | gtggaagaac  | ttggagacgc | ccttgtgacc | tccaagattt | 4320 |
|            |            | gatggcaatg  |            |            |            |      |
| tttctgggat | cactaacgtc | atagttgtgt  | tccaggatga | gatcgtcata | ggccattttt | 4440 |
| acaaagcgcg | ggcggagggt | gccagactgc  | ggtataatgg | ttccatccgg | cccaggggcg | 4500 |
| tagttaccct | cacagatttg | catttcccac  | gctttgagtt | cagatggggg | gatcatgtct | 4560 |
|            |            | aacggtttcc  |            |            |            |      |
|            |            | cttaccgcag  |            |            |            |      |
| ggctgcaact | ggtagttaag | agagetgeag  | ctgccgtcat | ccctgagcag | gggggccact | 4740 |
|            |            | tcgcatgttt  |            |            |            |      |
| ccgcccagcg | atagcagttc | ttgcaaggaa  | gcaaagtttt | tcaacggttt | gagaccgtcc | 4860 |
|            |            | cgtttgacca  |            |            |            |      |
|            |            | atccagcata  |            |            |            |      |
| cgctgtacgg | cagtagtcgg | tgctcgtcca  | gacgggccag | ggtcatgtct | ttccacgggc | 5040 |
|            |            | gtctgggtca  |            |            |            |      |
|            |            | ctggtcctgc  |            |            |            |      |
| gcgcgtcggc | caggtagcat | ttgaccatgg  | tgtcatagtc | cagcccctcc | gcggcgtggc | 5220 |
| ccttggcgcg | cagettgece | ttggaggagg  | cgccgcacga | ggggcagtgc | agacttttga | 5280 |
|            |            |             |            |            |            |      |

```
gggcgtagag cttgggcgcg agaaataccg attccgggga gtaggcatcc gcgccgcagg 5340
ccccgcagac ggtctcgcat tccacgagcc aggtgagctc tggccgttcg gggtcaaaaa 5400
ccaqqtttcc cccatgcttt ttgatgcgtt tcttacctct ggtttccatg agccggtgtc 5460
cacqctcggt gacgaaaagg ctgtccgtgt ccccgtatac agacttgaga ggcctgtcct 5520
cgagcggtgt tccgcggtcc tcctcgtata gaaactcgga ccactctgag acaaaggctc 5580
gcgtccaggc cagcacgaag gaggctaagt gggaggggta gcggtcgttg tccactaggg 5640
ggtccactcg ctccagggtg tgaagacaca tgtcgccctc ttcggcatca aggaaggtga 5700
ttggtttgta ggtgtaggcc acgtgaccgg gtgttcctga aggggggcta taaaaggggg 5760
tgggggcgcg ttcgtcctca ctctcttccg catcgctgtc tgcgagggcc agctgttggg 5820
gtgagtcgac gcgaggctgg atggccttcc ccattatgat tcttctcgct tccggcggca 5880
tegggatgee egegttgeag gecatgetgt ceaggeaggt agatgaegae cateagggae 5940
agetteaagg ccaqcaaaag gccaggaacc gtaaaaaaggc cgcgttgctg gcgtttttcc 6000
ataggeteeg ecceetgae gageateaea aaaategaeg eteaagteag aggtggegaa 6060
accogacage actataaaga taccaggogt ttccccctgg aagctccctc gtgcgctctc 6120
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 6180
cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 6240
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 6300
gtcttgagtc caaccoggta agacacgact tatcgccact ggcagcagcc actggtaaca 6360
qqattaqcaq aqcqaqqtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 6420
acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 6480
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 6540
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 6600
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 6660
qáttatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 6720
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 6780
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga 6840
taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 6900
cacgeteace ggetecagat ttateageaa taaaceagee ageeggaagg geegagegea 6960
quagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 7020
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct gcaggcatcg 7080
tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 7140
gagttacatg atoccccatg ttgtgcaaaa aagoggttag otocttoggt octoogatog 7200
ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 7260
ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 7320
cattctqaqa atagtgtatg eggegacega gttgetettg ceeggegtea acaegggata 7380
ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 7440
gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 7500
ccaactgate tteageatet tttactttea ccagegttte tgggtgagea aaaacaggaa 7560
ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 7620
tcctttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 7680
ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 7740
cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7800
                                                                   7828
cgaggcctt tcgtcttcaa gaattgtt
```

```
<210> 48
<211> 7789
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pAdApt containing HIV genes
<400> 48
```

| 1.1 1      |            | <b></b>    |            |            |            | <b>CO</b> |
|------------|------------|------------|------------|------------|------------|-----------|
|            | gcaattctca |            |            |            |            |           |
|            | caatatgata |            |            |            |            |           |
|            | tgacgtagta |            |            |            |            |           |
| tgtaagcgac | ggatgtggca | aaagtgacgt | ttttggtgtg | cgccggtgta | cacaggaagt | 240       |
| gacaattttc | gcgcggtttt | aggcggatgt | tgtagtaaat | ttgggcgtaa | ccgagtaaga | 300       |
|            | ttcgcgggaa |            |            |            |            |           |
|            | cgtaatattt |            |            |            |            |           |
| caccesages | tttttctcag | atattttca  | cattccaaat | caaagttggc | ottttattat | 480       |
| tatactcact | acgtaccagt | gcactggcct | agagegee   | cabbocabac | gttgtatcca | 540       |
| tatagecage | tgtacattta | tattogctca | totccaacat | taccgccatg | ttgacattga | 600       |
|            | gttattaata |            |            |            |            |           |
|            | ttacataact |            |            |            |            |           |
|            |            |            |            |            |            |           |
|            | cgtcaataat |            |            |            |            |           |
|            | gggtggagta |            |            |            |            |           |
|            | gtacgccccc |            |            |            |            |           |
|            | tgaccttatg |            |            |            |            |           |
|            | tggtgatgcg |            |            |            |            |           |
|            | ttccaagtct |            |            |            |            |           |
|            | actttccaaa |            |            |            |            |           |
|            | ggtgggaggt |            |            |            |            |           |
| tggagacgcc | atccacgctg | ttttgacctc | catagaagac | accgggaccg | atccagcctc | 1260      |
| cgtcaccgtc | gtcgacacgt | gtgatcagat | atcgcggccg | ctctagaatg | cgcgtgaagg | 1320      |
|            | gcacctgtgg |            |            |            |            |           |
|            | cagcgccacc |            |            |            |            |           |
|            | caccaccacc |            |            |            |            |           |
|            | gtgggccacc |            |            |            |            |           |
|            | cgtgaccgag |            |            |            |            |           |
|            | catcagectg |            |            |            |            |           |
|            | cctgaagtgc |            |            |            |            |           |
|            | catggagaag |            |            |            |            |           |
|            | gatcacccag |            |            |            |            |           |
|            |            |            |            |            |            |           |
|            | cgccggcttc |            |            |            |            |           |
|            | caacgtgagc |            |            |            |            |           |
|            | gctgaacggt |            |            |            |            |           |
|            | cgccaaggtg |            |            |            |            |           |
| cccgcgctag | cgcccactgc | aacctgagcc | gegeeaageg | gaacgacacc | ctgaacaaga | 2160      |
|            | gctgcgcgag |            |            |            |            |           |
|            | cgagatcgtg |            |            |            |            |           |
|            | cctgccctgc |            |            |            |            |           |
|            | cgccccccc  |            |            |            |            |           |
| tgctgctgac | ccgcgacggc | ggcaacagca | acaacgagag | cgagatette | cgtccgggcg | 2460      |
|            | gcgcgacaac |            |            |            |            |           |
|            | cgtggccccc |            |            |            |            |           |
|            | gcagcagcag |            |            |            |            |           |
| tgcagctgac | cgtgtggggc | atcaagcagc | tgcaggcccg | caccctggcc | gtggagcgct | 2,700     |
| acctgaagga | ccagcagcag | ctcgagcaga | tctggaacca | caccacctgg | atggagtggg | 2760      |
|            | caacaactac |            |            |            |            |           |
|            | gaacgagcag |            |            |            |            |           |
|            | caccaactgg |            |            |            |            |           |
|            | tgtttgcccc |            |            |            |            |           |
| ctatcettte | ctaataaaat | gaggaaattq | categeatte | tctgagtagg | tgtcattcta | 3060      |
|            | tggggtgggg |            |            |            |            |           |
|            | tgcggtgggc |            |            |            |            |           |
|            | gttataagat |            |            |            |            |           |
|            | gtgggggtct |            |            |            |            |           |
|            | J-33335000 |            |            |            |            |           |

| atgagcacca | actcqtttqa   | tggaagcatt | gtgagctcat  | atttgacaac | gcgcatgccc | 3360         |
|------------|--------------|------------|-------------|------------|------------|--------------|
| ccatgggccg | qqqtqcqtca   | gaatgtgatg | qqctccagca  | ttgatggtcg | ccccgtcctg | 3420         |
| cccqcaaact | ctactacctt   | gacctacgag | accgtgtctg  | gaacgccgtt | ggagactgca | 3480         |
| acctccacca | ccqcttcaqc   | cgctgcagcc | accgcccgcg  | ggattgtgac | tgactttgct | 3540         |
| ttcctgagcc | cacttacaaa   | cagtgcagct | tcccqttcat  | ccgcccgcga | tgacaagttg | 3600         |
| acqqctcttt | tggcacaatt   | ggattctttg | acccqqqaac  | ttaatgtcgt | ttctcagcag | 3660         |
| ctgttggatc | tacaccaaca   | aatttctacc | ctgaaggctt  | cctcccctcc | caatgcggtt | 3720         |
| taaaacataa | ataaaaaacc   | agactctgtt | tagatttaga  | tcaagcaagt | gtcttgctgt | 3780         |
| ctttatttag | gaattttaca   | cacacaataa | acccadascc  | agcggtctcg | gtcgttgagg | 3840         |
| atcetatata | tttttccag    | gacgtggtaa | aggtgactct  | ggatgttcag | atacatgggc | 3900         |
| ataagccgt  | ctctagagta   | gaggtagcac | cactgcagag  | cttcatqctq | cggggtggtg | 3960         |
| ttatagataa | tecagtegta   | gcaggagcgc | tagacataat  | gcctaaaaat | gtctttcagt | 4020         |
| agcaagctga | ttaccaaaaa   | caggeettg  | atataaatat  | ttacaaagcg | gttaagctgg | 4080         |
| ageaageega | tacataggaa   | tatgagatgc | atcttggact  | gtattttag  | gttggctatg | 4140         |
| ttccaacca  | tateceteeq   | gggattcatg | ttgtgcagaa  | ccaccaqcac | agtgtatccg | 4200         |
| atacacttaa | gaaatttgtc   | atgragetta | gaaggaaatg  | cotogaagaa | cttggagacg | 4260         |
| cccttatasc | ctccaagatt   | ttccatgcat | tcatccataa  | tgatggcaat | gggcccacgg | 4320         |
| acacacacat | aaacaaaaat   | atttctggga | tcactaacqt  | catagttoto | ttccaggatg | 4380         |
| agatogtoat | aggegaagae   | tacaaaacac | adacadaaaa  | taccagacta | cggtataatg | 4440         |
| agaccyccac | aggecaceee   | ctactaccc  | tcacagattt  | gcatttccca | cgctttgagt | 4500         |
| griceateeg | geeeaggge    | tacctacca  | acastasaas  | aaacootttc | cggggtaggg | 4560         |
| ceagatgggg | ggattatgtt   | cacctgcggg | accactaca   | acttacccca | gccggtgggc | 4620         |
| gagateaget | gggaagaaag   | caggereers | tactacttaa  | gagagetgea | gctgccgtca | 4680         |
| ccgcaaacca | cacctattac   | thochtonoo | atatacataa  | ctcacatatt | ttccctgacc | 4740         |
| ceeeegagea | ggggggccac   | cccgccaage | gatageagth  | cttgcaagga | agcaaagttt | 4800         |
| aaacccgcca | gaaggegete   | geegeeeage | atastatta   | acatttaaca | agcaaagttc | 4860         |
| ttcaacggtt | tgagaeegee   | egeegeagge | acgeeetete  | gegeeegaee | aagcagttcc | 4920         |
| aggeggteee | acageteggt   | baatataa   | acggcacccc  | gacccagcac | atctcctcgt | 4980         |
| ttegegggtt | ggggcggctt   | eggergracy | testeseest  | agtetegete | agacgggcca | 5040         |
| gggtcatgtc | tttccacggg   | cgcagggtcc | tegreagege  | agtetgggte | acggtgaagg | 5090         |
| garacacrec | gggctgcgcg   | ctggccaggg | cgegettgag  | thtasaasta | ctggtgctga | 5160         |
| agcgctgccg | gtettegeee   | tgegegtegg | ccaggragea  | attgaccacg | gtgtcatagt | 5220         |
| ccagcccctc | egeggegegg   | eccttggege | geagerrace  | gagagagag  | gegeegeacg | 5280         |
| aggggcagcg | cagacttttg   | agggcgtaga | gertgggege  | ttaaaaaaaa | gattccgggg | 5340         |
| agtaggcatc | egegeegeag   | geeeegeaga | aggeotegea  | tttaataat  | caggtgagct | 5400         |
| etggeegtte | ggggccaaaa   | accaggitte | tanganana   | actatacata | ttcttacctc | 5460         |
| tggtttecat | gageeggtgt   | teangeregg | thacacacata | ctcctcctat | teccegtata | 5520         |
| cagacttgag | aggeetgee    | ccgagcggtg | accagagaga  | ggagggtaag | agaaactcgg | 5580         |
| accactetga | gacaaaggct   | egegeedagg | ccagcacgaa  | ggaggccaag | tgggaggggt | 5640         |
| ageggtegtt | gtccactagg   | gggtecaete | geteeagggt  | grgaagacac | atgtcgccct | 5700         |
| cttcggcatc | aaggaaggcg   | attggtttgt | aggiglagge  | actotottoo | ggtgttcctg | 5760         |
| aagggggct  | acaaaagggg   | grgggggege | getegteete  | actecette  | gcatcgctgt | 5920         |
| ctgcgagggc | cagetgttgg   | ggtgagtega | egegaggerg  | gacggccccc | cccattatga | 5880         |
| ttettetege | teceggegge   | acceggacec | gggggggga   | ggccacgccg | tccaggcagg | 5940         |
| tagatgacga | ccatcaggga   | cagetteaag | gecageaaaa  | ggccaggaac | cgtaaaaagg | 5000         |
| ccgcgttgct | ggcgttttc    | eataggetee | geeeeeetga  | cyagoaccac | aaaaatcgac | 6060         |
| gctcaagtca | gaggtggcga   | aaccegacag | gactataaag  | acaccaggeg | tttccccctg | 6120         |
| gaageteeet | catacacter   | cccgccccga | eccegecgec  | taccygatac | ctgtccgcct | 6120         |
| ttctcccttc | gggaagcgtg   | gegettete  | atageteacg  | cigiaggiat | ctcagttcgg | 6240<br>6780 |
| tgtaggtcgt | regetecaag   | ccgggccgcg | Lgcacgaacc  | ceeegeeag  | cccgaccgct | 6240         |
| gcgccttato | : cggtaactat | cgtcttgagt | ccaacccggt  | aagacacgac | ttatcgccac | 6300         |
| tggcagcagc | cactggtaac   | aggattagca | gagcgaggta  | cgcaggcggt | gctacagagt | 0300         |
| tcttgaagtg | grggcctaac   | tacggctaca | cragaaggac  | agtatttggt | atctgcgctc | 6420         |
| tgctgaagco | agttaccttc   | ggaaaaagag | ctggtagete  | tegateegge | aaacaaacca | 6540         |
| ccgctggtag | ı cggtggtttt | tetgetegea | agcagcagac  | Lacycycaga | aaaaaaggat | 6600         |
| ctcaagaaga | tcctttgatc   | ttttctacgg | ggcccgacgc  | ccaycygaac | gaaaactcac | 0000         |

```
gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 6660
aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc 6720
aatgettaat cagtgaggea cetateteag egatetgtet atttegttea tecatagttg 6780
cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg 6840
ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc 6900
cagcoggaag ggccgagogc agaagtggto ctgcaacttt atccgcctcc atccagtcta 6960
ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg 7020
ttgccattgc tgcaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct 7080
ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta 7140
gctccttcqq tcctccqatc gttqtcagaa gtaaqttqqc cqcaqtqtta tcactcatqq 7200
ttatggcage actgcataat tetettactg teatgccate egtaagatge ttttetqtga 7260
ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt 7320
gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 7380
ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 7440
cgatgtaacc cactegtgca cccaactgat cttcagcatc ttttactttc accagegttt 7500
ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga 7560
aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat cagggttatt 7620
gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 7680
gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa 7740
cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattgtt
```

<210> 49 <211> 7711 <212> DNA <213> Artificial Sequence <220> <223> plasmid pAdApt containing HIV genes ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60 ggattgaago caatatgata atgagggggt ggagtttgtg acgtggcgcg gggcgtggga 120 acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180 tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240 gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300 tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360 acteatageg egtaatattt gtetagggee geggggaett tgacegttta egtggagaet 420 cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480 tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540 tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600 ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660 gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720 cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780 tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840 catatgccaa gtacgcccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900 geccagtaca tgacettatg ggaettteet acttggeagt acatetacgt attagteate 960 gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020 tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080 aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140 aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200 tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260 cgtcaccgtc gtcgacacgt gtgatcagat atcgcggccg ctctagaatg cgcgtgaagg 1320 agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380

|             |            |             |            |             |            | 2440 |
|-------------|------------|-------------|------------|-------------|------------|------|
|             |            |             |            |             | gtgcccgtgt |      |
|             |            |             |            |             | gacaccgagg |      |
| tgcacaacgt  | gtgggccacc | cacgcctgcg  | tgcccaccga | ccccaacccc  | caggaggtgg | 1560 |
|             |            |             |            |             | gagcagatgc |      |
| acgaggacat  | catcagcctg | tgggaccaga  | gcctgaagcc | ctgcgtgaag  | ctgacccccc | 1680 |
| tgtgcgtgag  | cctgaagtgc | accgacgcta  | gcaccagctg | caacaccagc  | gtgatcaccc | 1740 |
| aggeetgeee  | caaggtgagc | ttcgagccca  | tccccaacca | ctactgcgcc  | cccgccggct | 1800 |
| tcqccatcct  | gaagtgcaag | gacaagaagt  | tcaacggcaa | gggcccctgc  | accaacgtga | 1860 |
| gcaccgtgca  | gtgcacccac | ggcatccgcc  | ccgtggtgag | cacccagctg  | ctgctgaacg | 1920 |
|             |            |             |            |             | aacgccaagg |      |
|             |            |             |            |             | agcgcccact |      |
|             |            |             |            |             | aagctgcgcg |      |
|             |            |             |            |             | cccgagatcg |      |
| tgacccacag  | cttcaactgc | aacaacaaat  | tettetacto | caacaqcqct  | agcctgccct | 2220 |
| accocaticas | gragatrate | aacatotooc  | agaaggtggg | caaggccatg  | tacgcccccc | 2280 |
| ccatcaccaa  | ccadatccdc | tacaacaaca  | acatcaccog | cetactacta  | acccgcgacg | 2340 |
| ccaccagcgg  | caacaaccac | ancharatet  | tecatecaga | caacaacaac  | atgegegaca | 2400 |
| actggcaacag | caacaacgag | agogagaece  | taataaaat  | caaaccccta  | ggcgtggccc | 2460 |
| accegegeag  | caagettace | atccadaccc  | accaactact | gaggggatc   | gtgcagcagc | 2520 |
| ccaccaagge  | antacacaca | ategaggee   | accageeget | actacaacta  | accgtgtggg | 2580 |
| agaacaacct  | getgegegee | cacacacataa | ageageacee | ctacctgaag  | gaccagcagc | 2640 |
| gcatcaagca  | getgeaggee | caccecage   | ggatggagtg | ggaggag     | atcaacaact | 2700 |
| agelegagea  | gatteggaat | cacaccaccc  | 2940994909 | ggaccgcgag  | aagaacgagc | 2760 |
| acaccayeet  | gatteatage | ccgaccgagg  | agagecagaa | ctagtaggag  | atcaccaact | 2820 |
| aggagetget  | ggagerggae | tataatataa  | gcccgcggaa | ccagccatat  | attatttacc | 2880 |
| ggetgtggtg  | aggatecaga | congenginge | cttttagttg | ccagccaccc  | gttgtttgcc | 2000 |
| cctcccccgt  | geetteetty | accetggaag  | graceactec | tattatasa   | tcctaataaa | 2000 |
| acgaggaaac  | tgeategeat | tgtetgagta  | ggtgtcattc | cacteetgggg | ggtggggtgg | 3060 |
| ggcagcacag  | caagggggag | gattgggaag  | acaatageag | geargerggg  | gatgcggtgg | 3120 |
| getetatggg  | tacceaggge | cgcacaaccc  | cgtataatgt | acyclatacy  | aagttataag | 2100 |
| acctgtactg  | aaatgtgtgg | gegeggeeta  | agggtgggaa | agaacacaca  | aggtgggggt | 3240 |
| cttatgtagt  | tergratery | ttttgeagea  | geegeegeeg | ccatgageae  | caactcgttt | 3240 |
| gatggaagca  | ttgtgagete | acatttgaca  | acgegeatge | terearge    | cggggtgcgt | 3360 |
| cagaatgtga  | tgggctccag | cattgatggt  | egeeeegeee | tgeeegeaaa  | ctctactacc | 3430 |
| ttgacctacg  | agaccgtgtc | tggaacgccg  | ttggagactg | cageeteege  | cgccgcttca | 3420 |
| gccgctgcag  | ccaccgcccg | cgggattgtg  | actgactttg | cttteetgag  | cccgcttgca | 3400 |
|             |            |             |            |             | tttggcacaa |      |
| ttggattctt  | tgacccggga | acttaatgtc  | gtttctcagc | agetgttgga  | tctgcgccag | 3600 |
| caggtttctg  | ccctgaaggc | ttcctcccct  | cccaatgcgg | tttaaaacat  | aaataaaaaa | 3660 |
| ccagactctg  | tttggatttg | gatcaagcaa  | graterracr | gtetttattt  | aggggttttg | 3720 |
| cgcgcgcggt  | aggcccggga | ccagcggtct  | cggtcgttga | gggtcctgtg  | tatttttcc  | 3780 |
| aggacgtggt  | aaaggtgact | ctggatgttc  | agatacatgg | gcataagccc  | gtctctgggg | 3840 |
| tggaggtagc  | accactgcag | agcttcatgc  | racaaaaraa | tgttgtagat  | gatccagtcg | 3900 |
| tagcaggagc  | gctgggcgtg | gtgcctaaaa  | atgtetttea | gtagcaagct  | gattgccagg | 3960 |
| ggcaggccct  | tggtgtaagt | gtttacaaag  | cggttaagct | gggatgggtg  | catacgtggg | 4020 |
| gatatgagat  | gcatcttgga | ctgtatttt   | aggttggcta | tgttcccage  | catatecete | 4080 |
| cggggattca  | tgttgtgcag | aaccaccagc  | acagtgtate | cggtgcactt  | gggaaatttg | 4140 |
| tcatgtagct  | tagaaggaaa | tgcgtggaag  | aacttggaga | cgcccttgtg  | acctccaaga | 4200 |
| ttttccatgc  | attegteeat | aatgatggca  | atgggcccac | aaacaacaac  | ctgggcgaag | 4260 |
|             |            |             |            |             | ataggccatt |      |
|             |            |             |            |             | cggcccaggg |      |
| gcgtagttac  | cctcacagat | ttgcatttcc  | cacgetttga | gttcagatgg  | ggggatcatg | 4440 |
| tctacctgcg  | gggcgatgaa | gaaaacggtt  | tccggggtag | gggagatcag  | ctgggaagaa | 4500 |
| agcaggttcc  | tgagcagctg | cgacttaccg  | cagccggtgg | gcccgtaaat  | cacacctatt | 4560 |
| accggctgca  | actggtagtt | aagagagctg  | cagetgeegt | catecetgag  | caggggggcc | 4620 |
| acttcgttaa  | gcatgtccct | gactcgcatg  | ttttccctga | ccaaatccgc  | cagaaggcgc | 4680 |

|            |            |            |            |            | tttgagaccg |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | ccacageteg |      |
| gtcacctgct | ctacggcatc | tcgatccagc | atatetecte | gtttcgcggg | ttggggcggc | 4860 |
| tttcgctgta | cggcagtagt | cggtgctcgt | ccagacgggc | cagggtcatg | tctttccacg | 4920 |
| ggcgcagggt | cctcgtcagc | gtagtctggg | tcacggtgaa | ggggtgcgct | ccgggctgcg | 4980 |
| cgctggccag | ggtgcgcttg | aggctggtcc | tgctggtgct | gaagcgctgc | cggtcttcgc | 5040 |
| cctgcgcgtc | ggccaggtag | catttgacca | tggtgtcata | gtccagcccc | teegeggegt | 5100 |
| ggcccttggc | gcgcagcttg | cccttggagg | aggcgccgca | cgaggggcag | tgcagacttt | 5160 |
| tgagggcgta | gagcttgggc | gcgagaaata | ccgattccgg | ggagtaggca | teegegeege | 5220 |
| aggccccgca | gacggtctcg | cattccacga | gccaggtgag | ctctggccgt | tcggggtcaa | 5280 |
| aaaccaggtt | tececcatge | tttttgatgc | gtttcttacc | tctggtttcc | atgagccggt | 5340 |
|            |            |            |            |            | agaggcctgt |      |
|            |            |            |            |            | gagacaaagg |      |
| ctcgcgtcca | ggccagcacg | aaggaggcta | agtgggaggg | gtagcggtcg | ttgtccacta | 5520 |
| gggggtccac | tcgctccagg | gtgtgaagac | acatgtcgcc | ctcttcggca | tcaaggaagg | 5580 |
| tgattggttt | gtaggtgtag | gccacgtgac | cgggtgttcc | tgaagggggg | ctataaaagg | 5640 |
| gggtggggc  | gcgttcgtcc | tcactctctt | ccgcatcgct | gtctgcgagg | gccagctgtt | 5700 |
| ggggtgagtc | gacgcgaggc | tggatggcct | tccccattat | gattcttctc | gcttccggcg | 5760 |
| gcatcgggat | gcccgcgttg | caggccatgc | tgtccaggca | ggtagatgac | gaccatcagg | 5820 |
| gacagettea | aggccagcaa | aaggccagga | accgtaaaaa | ggccgcgttg | ctggcgtttt | 5880 |
| tccataggct | ccgccccct  | gacgagcatc | acaaaaatcg | acgctcaagt | cagaggtggc | 5940 |
| gaaacccgac | aggactataa | agataccagg | cgtttccccc | tggaagctcc | ctcgtgcgct | 6000 |
| ctcctqttcc | gaccctgccg | cttaccggat | acctgtccgc | ctttctccct | tcgggaagcg | 6060 |
| tggcgctttc | tcatagetca | cgctgtaggt | atctcagttc | ggtgtaggtc | gttcgctcca | 6120 |
| agetgggetg | tgtgcacgaa | cccccgttc  | agcccgaccg | ctgcgcctta | tccggtaact | 6180 |
| atcqtcttga | gtccaacccg | gtaagacacg | acttatcgcc | actggcagca | gccactggta | 6240 |
| acaggattag | cagagcgagg | tatgtaggcg | gtgctacaga | gttcttgaag | tggtggccta | 6300 |
| actacggcta | cactagaagg | acagtatttg | gtatctgcgc | tctgctgaag | ccagttacct | 6360 |
|            |            |            |            |            | agcggtggtt |      |
|            |            |            |            |            | gatcctttga |      |
| tcttttctac | ggggtctgac | gctcagtgga | acgaaaactc | acgttaaggg | attttggtca | 6540 |
| tgagattatc | aaaaaggatc | ttcacctaga | tccttttaaa | ttaaaaatga | agttttaaat | 6600 |
| caatctaaag | tatatatgag | taaacttggt | ctgacagtta | ccaatgctta | atcagtgagg | 6660 |
| cacctatctc | agcgatctgt | ctatttcgtt | catccatagt | tgcctgactc | cccgtcgtgt | 6720 |
| agataactac | gatacgggag | ggcttaccat | ctggccccag | tgctgcaatg | ataccgcgag | 6780 |
|            |            |            |            |            | agggccgagc |      |
|            |            |            |            |            | tgccgggaag |      |
|            |            |            |            |            | gctgcaggca |      |
|            |            |            |            |            | caacgatcaa |      |
|            |            |            |            |            | ggtcctccga |      |
|            |            |            |            |            | gcactgcata |      |
|            |            |            |            |            | tactcaacca |      |
|            |            |            |            |            | tcaacacggg |      |
|            |            |            |            |            | cgttcttcgg |      |
|            |            |            |            |            | cccactcgtg |      |
|            |            |            |            |            | gcaaaaacag |      |
|            |            |            |            |            | atactcatac |      |
|            |            |            |            |            | agcggataca |      |
|            |            |            |            |            | ccccgaaaag |      |
|            |            |            |            | aacctataaa | aataggcgta |      |
| tcacgaggcc | ctttcgtctt | caagaattgt | t          |            |            | 7711 |

<211> 6577 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 50 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttqqcqqqtq tcqqqqctqq cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtacege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getataetgt tittggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattgqtqac qatactttcc attactaatc cataacatqq ctctttqcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg ageeetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggeeag acttaggeac 1560 agcacaatge ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag gctagcagcc tgaagtgcac cgacctgaag 2040 aacgacacca acaccaacag cagcagcggc cgcatgatca tggagaaggg cgagatcaag 2100 aactgcagct tcaacatcag caccagcatc cgcggcaagg tgcagaagga gtacgccttc 2160 ttctacaagc tggacatcat ccccatcgac aacgacacca ccagctacag cctgaccagc 2220 tgcaacacca gegtgatcac ccaggeetge cecaaggtga gettegagee cateeccaae 2280 cactactgcg cccccgccgg cttcgccatc ctgaagtgca aggacaagaa gttcaacggc 2340 aagggcccct gcaccaacgt gagcaccgtg cagtgcaccc acggcatccg ccccgtggtg 2400

agcacccage tgetgetgaa eggtageetg geegaggagg aggtggtgat eegeageget 2460 aacttegeeg acaacgeeaa ggtgateate gtgeagetga acgagaget ggagateaac 2520 tgeaccege ecaacaacaa eaceegeaag agcateeaca teggeeceegg eegegeette 2580 tacaccaceg gegagateat eggegacate egeeaggeee actgeaacet gageegegee 2640 aagtggaacg acaccetgaa eaagategtg ateaagetge gegageagtt eggeaacaag 2700 accategtgt teaageacag eageggege gacceegaga tegtgaecea eagetteaac 2760 tgeggeggeg agttetteta etgeaacage acceagetgt teaacageac etggtteaac 2820

|              |            | ~~~~~~~         | 2000000000 | anancecat  | carretgere | 2880 |
|--------------|------------|-----------------|------------|------------|------------|------|
| agcacctgga g | geacegaygg | cagcaacaac      | accgagggca | gegacaccac | atacacacac | 2000 |
| tgccgcatca a | agcagatcat | caacatgtgg      | cagaaggrgg | gcaaggccat | gcacgccccc | 2000 |
| cccatcagcg o | gccagatccg | crgcageage      | aacaccaccg | geetgetget | gaccegegac | 3000 |
| ggcggcaaca   | gcaacaacga | gagegagate      | tteegteegg | geggeggega | catgegegae | 3120 |
| aactggcgca 🤉 | gcgagctgta | caagtacaag      | grggrgaaga | tegageceet | gggegtggee | 3120 |
| cccaccaagg   | ccaagcttac | cgtccaggcc      | cgccagctgc | tgageggeat | egtgeageag | 3180 |
| cagaacaacc   | tgctgcgcgc | catcgaggcc      | cagcagcacc | tgctgcagct | gaccgtgtgg | 3240 |
| ggcatcaagc   | agctgcaggc | ccgcaccctg      | gccgtggagc | gctacctgaa | ggaccagcag | 3300 |
| cagctcgagc   | agatctggaa | ccacaccacc      | tggatggagt | gggaccgcga | gatcaacaac | 3360 |
| tacaccagcc   | tgatccacag | cctgatcgag      | gagagccaga | accagcacga | gaagaacgag | 3420 |
| caggagctgc   | tggagctgga | caagtgggcc      | agcctgtgga | actggttcaa | catcaccaac | 3480 |
| tggctgtggt   | acatcaagct | gttcatcatg      | atcgtgggcg | gcctggtggg | cctgcgcatc | 3540 |
| gtgttcgccg   | tgctgagcat | ctgaggatcc      | agatctgctg | tgccttctag | ttgccagcca | 3600 |
| tctgttgttt   | gecectecee | cgtgccttcc      | ttgaccctgg | aaggtgccac | tcccactgtc | 3660 |
| ctttcctaat   | aaaatgagga | aattgcatcg      | cattgtctga | gtaggtgtca | ttctattctg | 3720 |
| gggggtgggg   | tggggcagga | cagcaagggg      | gaggattggg | aagacaatag | caggcatgct | 3780 |
| ggggatgcgg   | tgggctctat | gggtacccag      | gtgctgaaga | attgacccgg | ttcctcctgg | 3840 |
| gccagaaaga   | agcaggcaca | teceettete      | tgtgacacac | cctgtccacg | cccctggttc | 3900 |
| ttagttccag   | cccactcat  | aggacactca      | tagctcagga | gggctccgcc | ttcaatccca | 3960 |
| cccgctaaag   | tacttggagc | ggtctctccc      | tccctcatca | gcccaccaaa | ccaaacctag | 4020 |
| cctccaagag   | tgggaagaaa | ttaaagcaag      | ataggctatt | aagtgcagag | ggagagaaaa | 4080 |
| tgcctccaac   | atgtgaggaa | gtaatgagag      | aaatcataga | attttaaggc | catcatggcc | 4140 |
| ttaatcttcc   | gcttcctcgc | tcactgactc      | gctgcgctcg | gtcgttcggc | tgcggcgagc | 4200 |
| ggtatcagct   | cactcaaagg | cggtaatacg      | gttatccaca | gaatcagggg | ataacgcagg | 4260 |
| aaagaacatg   | tgagcaaaag | gccagcaaaa      | ggccaggaac | cgtaaaaagg | ccgcgttgct | 4320 |
| gacatttttc   | cataggetee | gcccccctga      | cgagcatcac | aaaaatcgac | gctcaagtca | 4380 |
| gaggtggcga   | aacccgacag | gactataaag      | ataccaggcg | tttccccctg | gaagctccct | 4440 |
| catacactct   | cctgttccga | ccctgccgct      | taccggatac | ctgtccgcct | ttctcccttc | 4500 |
| gggaagcgtg   | gcgctttctc | atagctcacg      | ctgtaggtat | ctcagttcgg | tgtaggtcgt | 4560 |
| tcqctccaaq   | ctaggetgtg | tgcacgaacc      | ccccgttcag | cccgaccgct | gcgccttatc | 4620 |
| cogtaactat   | cgtcttgagt | ccaacccggt      | aagacacgac | ttatcgccac | tggcagcagc | 4680 |
| cactggtaac   | aggattagca | gagcgaggta      | tgtaggcggt | gctacagagt | tcttgaagtg | 4740 |
| gtggcctaac   | tacqqctaca | ctagaagaac      | agtatttggt | atctgcgctc | tgctgaagcc | 4800 |
| agttaccttc   | qqaaaaaqaq | ttggtagete      | ttgatccggc | aaacaaacca | ccgctggtag | 4860 |
| caataatttt   | tttqtttqca | agcagcagat      | tacgcgcaga | aaaaaaggat | ctcaagaaga | 4920 |
| tcctttgatc   | ttttctacqq | ggtctgacgc      | tcagtggaac | gaaaactcac | gttaagggat | 4980 |
| tttggtcatg   | agattatcaa | aaaqqatctt      | cacctagatc | cttttaaatt | aaaaatgaag | 5040 |
| tttaaatca    | atctaaagta | tatatgagta      | aacttggtct | gacagttacc | aatgcttaat | 5100 |
| cagtgaggca   | cctatctcag | cgatctgtct      | atttcqttca | tocatagttg | cctgactcgg | 5160 |
| aaaaaaaaaa   | cactaaaatc | tacctcataa      | agaaggtgtt | gctgactcat | accaggcctg | 5220 |
| aatcgccca    | tcatccagcc | agaaagtgag      | qqaqccacqq | ttgatgagag | ctttgttgta | 5280 |
|              |            |                 |            |            | cgttgtcggg |      |
| aagatgcgtg   | atctgatcct | tcaactcagc      | aaaagttcga | tttattcaac | aaagccgccg | 5400 |
| tcccatcaaa   | tcagcgtaat | getetgecag      | tottacaacc | aattaaccaa | ttctgattag | 5460 |
| aaaaactcat   | cgagcatcaa | atgaaactgc      | aatttattca | tatcaggatt | atcaatacca | 5520 |
| tatttttgaa   | aaagccgttt | ctgtaatgaa      | ggagaaaact | caccgaggca | gttccatagg | 5580 |
| atoocaagat   | cctootatco | atctacaatt      | ccgactcgtc | caacatcaat | acaacctatt | 5640 |
| aatttcccct   | catcaaaaat | aaggttatca      | agtgagaaat | caccatgagt | gacgactgaa | 5700 |
| tecaatasas   | atggcaaaag | cttatqcatt      | tctttccaqa | cttgttcaac | aggccagcca | 5760 |
| ttacgctcgt   | catcanaatc | actcocatca      | accaaaccot | tattcattca | tgattgcgcc | 5820 |
| taaacaaaac   | gaaatacgcg | atcoctotta      | aaaqqacaat | tacaaacaaa | aatcgaatgc | 5880 |
| aaccaacaca   | ggaacactgc | cagcgcatca      | acaatattt  | cacctgaatc | aggatattct | 5940 |
| totaatacct   | ggaatactat | tttcccaaaa      | atcqcaqtqq | tgagtaacca | tgcatcatca | 6000 |
| ggagtacgga   | taaaatactt | gatgatcaaa      | agaggcataa | attccatcaa | ccagtttagt | 6060 |
| ctgaccatct   | catchgtaac | atcattogca      | acqctacctt | toccatottt | cagaaacaac | 6120 |
| Legaceace    | Jacabagaaa | ~ <del>~-</del> | ,          | . 5        | J          |      |

```
tetggegeat egggettece atacaatega tagattgteg cacetgattg eeegacatta 6180 tegegagece atttatace atacaatea geatecatgt tggaatttaa tegeggeete 6240 gageaagacg tttecegttg aatatggete ataacacee ttgtattact gtttatgtaa 6300 geagacagtt ttattgtea tgatgatata tttttatett gtgeaatgta acateagaga 6360 ttttgagaca caaegtgget ttececece eeecattatt gaageattta teagggttat 6420 tgteteatga geggatacat atttgaatgt atttagaaaa ataaacaaat aggggtteeg 6480 eeecatataaaa ataggegtat eaegaggee tttegte 5577
```

<210> 51 <211> 6577 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes

tegegeetht eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtacege etatagaete tataggeaea eccetttgge 1080 tottatqcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac totgtcottc agagactgac acggactctg tatttttaca 1320 qqatqqqqtc ccatttatta tttacaaatt cacatataca acaacqccqt cccccqtqcc 1380 cqcaqttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 qcaqaaqaaq atqcaqqcaq ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cqcqccacca qacataataq ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tqcaqtcacc qtcqtcqaca cqtqtgatca gatatcgcgg ccgctctaga caccatgcgc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040 cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100

| accasaatac | acaacatata | ggccacccac   | acctacatac   | ccaccgaccc | caacccccaq | 2160 |
|------------|------------|--------------|--------------|------------|------------|------|
| gangtootoc | taataaacat | gaccgagaac   | ttcgacatgt   | ggaagaacga | catqqtqqaq | 2220 |
| cacatocaco | aggagatest | cacctataa    | gaccagagcc   | tgaagccctg | cataaaacta | 2280 |
| accecetat  | acatasacct | gaagtgcacc   | gacctgaaga   | acgacaccaa | caccaacaqc | 2340 |
| acceceege  | gcgtgagcet | adadaaddac   | gagatcaaga   | actgcagctt | caacatcagc | 2400 |
| ageageggee | gcacgaccac | deadardade   | tacaccttct   | tctacaagct | ggacatcatc | 2460 |
| accagcaccc | 200300000  | geagaaggag   | chaccaact    | gcaacaccgc | tagcatcaac | 2520 |
| terrange   | acqueaccac | cagecacage   | aggatogge    | teggeceegg | ccacacatta | 2580 |
| tgcacccgcc | ccaacaacaa | caccegeaag   | caccacaca    | actgcaacct | asaccacaca | 2640 |
| cacaccaccg | gegagateat | cooccatcata  | atcaagetge   | gcgagcagtt | caacaacaaa | 2700 |
| aagtggaacg | teacectgaa | caagaccgcg   | gaccccaaga   | togtgaccca | cagetheaac | 2760 |
| accategigi | ccaagcacag | cageggegge   | accecgaga    | tcgtgaccca | ctoottcaac | 2820 |
| cgcggcggcg | agttetteta | ergeaacage   | acceageege   | tcaacagcac | caccetacce | 2880 |
| agcacctgga | gcaccgaggg | eagcaacaac   | accyagggca   | gcgacaccat | atacaccac  | 2940 |
| tgccgcatca | agcagaccac | caacatgtgg   | cagaaggcgg   | gcaaggccat | geacgeecee | 3000 |
| cccatcagcg | gecagateeg | etgeageage   | addattaccg   | gcctgctgct | gattegegat | 3060 |
| ggcggcaaca | gcaacaacga | gagegagate   | ctccgtccgg   | gcggcggcga | cacgegegae | 3120 |
| aactggcgca | gcgagctgta | caagtacaag   | gtggtgaaga   | tcgagcccct | gggcgcggcc | 2190 |
| cccaccaagg | ccaagcttac | cgtccaggcc   | egecagetge   | tgagcggcat | cgcgcagcag | 3700 |
| cagaacaacc | tgctgcgcgc | catcgaggcc   | cagcagcacc   | tgctgcagct | gacegegegg | 3240 |
| ggcatcaagc | agctgcaggc | ccgcaccctg   | gccgrggagc   | gctacctgaa | ggaccagcag | 3300 |
| cagctcgagc | agatctggaa | ccacaccacc   | tggatggagt   | gggaccgcga | gatcaacaac | 3360 |
| tacaccagec | tgatccacag | cctgatcgag   | gagagccaga   | accagcacga | gaagaacgag | 3420 |
| caggagctgc | tggagctgga | caagtgggcc   | agcctgtgga   | actggttcaa | catcaccaac | 3480 |
| tggctgtggt | acatcaagct | gttcatcatg   | atcgtgggcg   | gcctggtggg | cctgcgcatc | 3540 |
| gtgttcgccg | tgctgagcat | ctgaggatcc   | agatctgctg   | tgccttctag | ttgccagcca | 3600 |
| tctgttgttt | gcccctcccc | cgtgccttcc   | ttgaccctgg   | aaggtgccac | teccaetgte | 3660 |
| ctttcctaat | aaaatgagga | aattgcatcg   | cattgtctga   | gtaggtgtca | ttctattctg | 3720 |
| gggggtgggg | tggggcagga | cagcaagggg   | gaggattggg   | aagacaatag | caggcatgct | 3780 |
| ggggatgcgg | tgggctctat | gggtacccag   | gtgctgaaga   | attgacccgg | ttectectgg | 3840 |
| gccagaaaga | agcaggcaca | teceettete   | tgtgacacac   | cctgtccacg | cccctggttc | 3900 |
| ttagttccag | ccccactcat | aggacactca   | tagctcagga   | gggctccgcc | ttcaatccca | 3960 |
| cccgctaaag | tacttggagc | ggtctctccc   | tccctcatca   | gcccaccaaa | ccaaacctag | 4020 |
| cctccaagag | tgggaagaaa | ttaaagcaag   | ataggctatt   | aagtgcagag | ggagagaaaa | 4080 |
| tgcctccaac | atgtgaggaa | gtaatgagag   | aaatcataga   | attttaaggc | catcatggcc | 4140 |
| ttaatcttcc | gcttcctcgc | tcactgactc   | gctgcgctcg   | gtcgttcggc | tgcggcgagc | 4200 |
| ggtatcagct | cactcaaagg | cggtaatacg   | gttatccaca   | gaatcagggg | ataacgcagg | 4260 |
| aaagaacatg | tgagcaaaag | gccagcaaaa   | ggccaggaac   | cgtaaaaagg | ccgcgttgct | 4320 |
| ggcgtttttc | cataggctcc | gccccctga    | cgagcatcac   | aaaaatcgac | gctcaagtca | 4380 |
| qaggtggcga | aacccgacag | gactataaag   | ataccaggcg   | tttccccctg | gaagctccct | 4440 |
| catacactet | cctgttccga | ccctgccgct   | taccggatac   | ctgtccgcct | ttctcccttc | 4500 |
| gggaagcgtg | gcgctttctc | atagctcacg   | ctgtaggtat   | ctcagttcgg | tgtaggtcgt | 4560 |
| tegetecaag | ctgggctgtg | tgcacgaacc   | ccccgttcag   | cccgaccgct | gcgccttatc | 4620 |
| cggtaactat | cgtcttgagt | ccaacccggt   | aagacacgac   | ttatcgccac | tggcagcagc | 4680 |
| cactggtaac | aggattagca | gagcgaggta   | tgtaggcggt   | gctacagagt | tcttgaagtg | 4740 |
| gtggcctaac | tacggctaca | ctagaagaac   | agtatttggt   | atctgcgctc | tgctgaagcc | 4800 |
| agttaccttc | qqaaaaaqaq | ttggtagctc   | ttgatccggc   | aaacaaacca | ccgctggtag | 4860 |
| caataatttt | tttqtttqca | agcagcagat   | tacgcgcaga   | aaaaaaggat | ctcaagaaga | 4920 |
| teetttgate | ttttctacqq | ggtctgacgc   | tcagtggaac   | gaaaactcac | gttaagggat | 4980 |
| tttggtcatg | agattatcaa | aaaggatctt   | cacctagate   | cttttaaatt | aaaaatgaag | 5040 |
| ttttaaatca | atctaaagta | tatatgagta   | aacttagtct   | gacagttacc | aatgcttaat | 5100 |
| cagtgaggca | cctatctcac | cgatctgtct   | atttcqttca   | tccatagttg | cctgactcgg | 5160 |
| 0000000000 | cactaeaata | tacctcata    | agaaggtgtt   | gctgactcat | accaggeetq | 5220 |
| aatcoccce  | tcatccagc  | : agaaagtgag | ggagccacqq   | ttgatgagag | ctttgttgta | 5280 |
| aataaaccaa | ttagtaattt | : tgaacttttc | ctttqccacq   | gaacggtctg | cgttgtcggg | 5340 |
| aagatacata | atctgatect | tcaactcaoc   | : aaaaqttcqa | tttattcaac | aaagccgccq | 5400 |
| ~~~~~~~~   |            |              |              |            |            |      |

```
tecegteaag teagegtaat getetgeeag tgttacaace aattaaceaa ttetgattag 5460
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5520
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5580
atggcaagat cctggtateg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5640
aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5700
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5760
ttacgctcqt catcaaaatc actcqcatca accaaaccqt tattcattcq tgattqcqcc 5820
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 5880
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 5940
tetaatacet ggaatgetgt tttcccgggg atcgcagtgg tgagtaacca tgcatcatca 6000
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6060
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6120
tetggegeat egggetteee atacaatega tagattgteg cacetgattg ceegacatta 6180
tegegagece atttatacee atataaatea geateeatgt tggaatttaa tegeggeete 6240
gagcaagacg threecottg aatatggete ataacaccce tigtattact gittatgtaa 6300
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6360
ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta tcagggttat 6420
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6480
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6540
acctataaaa ataggcgtat cacgaggccc tttcgtc
                                                                  6577
```

<210> 52

<211> 6718 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teageggege teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattqqcca ttqcatacqt tqtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatege tattaccatg gtgatgeggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg.cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac totgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380

| WO 02/3294 | 3          |            |            |            | PCT/US     | 501/257 |
|------------|------------|------------|------------|------------|------------|---------|
| cgcagtttt  | attaaacata | gcgtgggatc | tccacgcgaa | tctcgggtac | gtgttccqqa | 1440    |
| catgggctct | tctccggtag | cggcggagct | tccacatccg | agccctggtc | ccatgcctcc | 1500    |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tggaggccag | acttaggcac | 1560    |
| agcacaatgo | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620    |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact | taaqqcaqcq | 1680    |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740    |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tactaccaca | 1800    |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | tcctttccat | gggtcttttc | 1860    |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatategegg | ccgctctaga | caccatqcqc | 1920    |
| gtgaaggaga | agtaccagca | cctgtggcgc | tggggctggc | getggggcae | catactacta | 1980    |
| ggcatgctga | tgatctgcag | cgccaccgag | aagctgtggg | tgaccgtgta | ctacggcgtg | 2040    |
| cccgtgtgga | aggaggccac | caccaccctg | ctctgcgcca | gcgacgccaa | ggcctacgac | 2100    |
| accgaggtgc | acaacgtgtg | ggccacccac | gcctgcgtgc | ccaccqaccc | caacccccag | 2160    |
| gaggtggtgc | tggtgaacgt | gaccgagaac | ttcgacatgt | qqaaqaacqa | catgotogag | 2220    |
| cagatgcacg | aggacatcat | cagcctgtgg | gaccagagee | tgaagccctg | catgaagcto | 2280    |
| accccctgt  | gcgtgagcct | gaagtgcacc | gacctgaaga | acgacaccaa | Caccaacage | 2340    |
| agcagcggcc | gcatgatcat | ggagaagggc | gagatcaaga | actgcagett | caacatcagc | 2400    |
| accagcatcc | gcggcaaggt | gcagaaggag | tacgccttct | tctacaaget | ggacatcatc | 2460    |
| cccatcgaca | acgacaccac | cagctacage | ctgaccaget | gcaacaccag | cataatcacc | 2520    |
| caggectgee | ccaaggtgag | cttcgagccc | atccccaacc | actactococ | cccaccaac  | 2580    |
| ttcgccatcc | tgaagtgcaa | ggacaagaag | ttcaacggca | agggccctg  | caccaacgtg | 2640    |
| agcaccgtgc | agtgcaccca | caacatccac | cccataataa | gcacccaget | actactasac | 2700    |
| ggtagcctgg | ccgaggagga | ggtggtgatc | cgcagcgcta | acttcgccga | caacqccaaq | 2760    |
| gtgatcatcg | tgcagctgaa | cqaqaqcqtq | gagatcaact | gcacccgccc | caacaacaac | 2820    |
| accegeaaqa | gcatccacat | caaccccaac | cacaccttct | acaccaccoo | cgagatcatc | 2880    |
| ggcgacatcc | gccaggccca | ctgcaacctg | gctagcttct | actgcaacag | cacccaacta | 2940    |
| ttcaacagca | cctggttcaa | cagcacctgg | agcaccgagg | gcagcaacaa | caccaaaaac | 3000    |
| agcgacacca | tcaccctgcc | ctgccgcatc | aagcagatca | tcaacatata | acagaaggta | 3060    |
| ggcaaggcca | tgtacgcccc | ccccatcage | ggccagatcc | gctgcagcag | caacatcacc | 3120    |
| ggcctgctgc | tgacccgcga | cggcggcaac | agcaacaacg | agagcgagat | cttccatcca | 3180    |
| ggcggcggcg | acatgcgcga | caactggcgc | agcgagctgt | acaagtacaa | aataataaaa | 3240    |
| atcgagcccc | tgggcgtggc | ccccaccaaq | gccaagetta | ccatccagac | ccaccaacta | 3300    |
| ctgagcggca | tcgtgcagca | gcagaacaac | ctactacaca | ccatcgaggc | ccagcagcac | 3360    |
| ctgctgcagc | tgaccgtgtg | gggcatcaaq | caqctqcaqq | cccgcaccct | gaccataasa | 3420    |
| cgctacctga | aggaccagca | gcagctcgag | cagatetgga | accacaccac | ctagatagaa | 3480    |
| tgggaccgcg | agatcaacaa | ctacaccage | ctgatccaca | gcctgatcga | ggagagccag | 3540    |
| aaccagcacg | agaagaacga | gcaggagctg | ctagaactag | acaagtgggc | cagcototog | 3600    |
| aactggttca | acatcaccaa | ctaactataa | tacatcaage | tgttcatcat | gatcgtggg  | 3660    |
| ggcctggtgg | gcctgcgcat | cgtqttcqcc | gtgctgagca | tctgaggatc | cagatetget | 3720    |
| gtgccttcta | gttgccagcc | atctgttgtt | tgcccctccc | ccataccttc | cttgaccctg | 3780    |
| gaaggtgcca | ctcccactgt | cctttcctaa | taaaatqaqq | aaattqcatc | gcattgtctg | 3840    |
| agtaggtgtc | attctattct | ggggggtggg | gtggggcagg | acagcaaggg | qqaqqattqq | 3900    |
| gaagacaata | gcaggcatgc | tggggatgcg | gtgggctcta | tqqqtaccca | ggtgctgaag | 3960    |
| aattgacccg | gttcctcctg | ggccagaaag | aagcaggcac | atccccttct | ctqtqacaca | 4020    |
| ccctgtccac | gcccctggtt | cttagttcca | gccccactca | taggacactc | atagctcagg | 4080    |
| agggctccgc | cttcaatccc | acccgctaaa | gtacttggag | caatctctcc | ctccctcatc | 4140    |
| agcccaccaa | accaaaccta | gcctccaaga | gtgggaagaa | attaaaqcaa | gataggetat | 4200    |
| taagtgcaga | gggagagaaa | atgcctccaa | catgtgagga | agtaatgaga | gaaatcatag | 4260    |
| aattttaagg | ccatcatggc | cttaatcttc | cacttcctca | ctcactgact | cactacacte | 4320    |
| ggtcgttcgg | ctgcggcgag | cggtatcagc | tcactcaaaa | gcggtaatac | aghtatorac | 4380    |
| agaatcaggg | gataacgcag | gaaagaacat | gtgagcaaaa | ggccagcaaa | addccaddaa | 4440    |
| ccgtaaaaag | gccgcgttgc | tggcqttttt | ccataggete | cacceceta  | acqaqcatca | 4500    |
| caaaaatcga | cgctcaagtc | agaggtggcg | aaacccqaca | ggactataaa | gataccagge | 4560    |
| gtttccccct | ggaagctccc | tegtacacte | tectatteca | accetaceae | ttaccoosts | 4620    |
| cctgtccgcc | tttctccctt | caggaagcat | ggcgctttct | catageteac | actatagata | 4680    |
|            |            | 5555-50    |            |            | arraraggia | 4000    |

```
teteagtteg gtgtaggteg ttegeteeaa getgggetgt gtgeaegaac ceceegttea 4740
qcccqaccqc tqcqccttat ccqqtaacta tcqtcttgag tccaacccgg taagacacga 4800
cttategeca etggeageag ccaetggtaa caggattage agagegaggt atgtaggegg 4860
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4920
tatctqcqct ctqctqaaqc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4980
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 5040
aaaaaaagga totcaagaag atootttgat ottttotacg gggtotgacg otcagtggaa 5100
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 5160
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5220
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 5280
atccatagtt gcctgactcg ggggggggg gcgctgaggt ctgcctcgtg aagaaggtgt 5340
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5400
qttqatqaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5460
ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag caaaagttcg 5520
atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5580
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttattc 5640
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5700
teacegagge agtteeatag gatggeaaga teetggtate ggtetgegat teegaetegt 5760
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5820
teaceatgag tgaegaetga ateeggtgag aatggeaaaa gettatgeat ttettteeag 5880
acttgttcaa caggocagoc attacgotog toatcaaaat cactogoato aaccaaacog 5940
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 6000
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 6060
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg gatcgcagtg 6120
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6180
aatteegtea geeagtttag tetgaceate teatetgtaa cateattgge aaegetaeet 6240
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6300
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6360
ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6420
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6480
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttcccccc cccccattat 6540
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6600
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6660
accattatta tcatgacatt aacctataaa aataggegta tcacgaggee etttegte
```

```
<211> 6742
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 53
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagcgggtg 120
ttggcgggtg tcggggctg cttaactatg cggcatcaga gcagattgta ctgagagtg 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacattata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
```

<210> 53

| tgcccacttg | gcagtacatc | aagtgtatca   | tatgccaagt | acgcccccta | ttgacgtcaa  | 600   |
|------------|------------|--------------|------------|------------|-------------|-------|
| tgacggtaaa | tggcccgcct | ggcattatgc   | ccagtacatg | accttatggg | actttcctac  | 660   |
| ttggcagtac | atctacgtat | tagtcatcgc   | tattaccatg | gtgatgcggt | tttggcagta  | 720   |
| catcaatggg | cgtggatagc | ggtttgactc   | acggggattt | ccaagtctcc | accccattga  | 780   |
| catcaataga | agtttgttt  | qqcaccaaaa   | tcaacgggac | tttccaaaat | gtcgtaacaa  | 840   |
| ctccacccca | ttgacgcaaa | tagacaataa   | gcgtgtacgg | tgggaggtct | atataagcag  | 900   |
| agetegttta | gtgaaccgtc | agatogootg   | gagacgccat | ccacactatt | ttgacctcca  | 960   |
| tagaagagag | cadaaccaat | ccagcctccg   | cggccgggaa | cootocatto | gaacgcggat  | 1020  |
| tecceatace | aagaggggg  | taantaccoc   | ctatagactc | tataggcaca | cccctttaac  | 1080  |
| tattatagat | actatactat | ttttaactta   | gggcctatac | acccccactt | ccttatgcta  | 1140  |
| tacatastaa | tatacettac | cctatagata   | tgggttattg | accettatta | accactocco  | 1200  |
| taggtgatgg | catagettag | cctataggtg   | catacatas  | statttagg  | gaaghatgtg  | 1260  |
| cattggtgae | gatacttcc  | actactaatt   | cataacatgg | ccccccgcca | tattttta    | 1220  |
| tattggctat | argccaatac | tetgteette   | agagactgac | acggactetg | catttttaca  | 1200  |
| ggatggggtc | ccatttatta | tttacaaatt   | cacatataca | acaacgccgt | eccegegee   | 1360  |
| cgcagttttt | attaaacata | gcgtgggatc   | tccacgcgaa | teregggrae | grgttccgga  | 1440  |
| catgggctct | tctccggtag | cggcggagct   | tccacatccg | agccctggtc | ccatgcctcc  | 1500  |
| agcggctcat | ggtcgctcgg | cagctccttg   | ctcctaacag | tggaggccag | acttaggcac  | 1560  |
| agcacaatgc | ccaccaccac | cagtgtgccg   | cacaaggccg | tggcggtagg | gtatgtgtct  | 1620  |
| gaaaatgagc | gtggagattg | ggctcgcacg   | gctgacgcag | atggaagact | taaggcagcg  | 1680  |
| gcagaagaag | atgcaggcag | ctgagttgtt   | gtattctgat | aagagtcaga | ggtaactccc  | 1740. |
| gttgcggtgc | tgttaacggt | ggagggcagt   | gtagtctgag | cagtactcgt | tgctgccgcg  | 1800  |
| cgcgccacca | gacataatag | ctgacagact   | aacagactgt | tcctttccat | gggtcttttc  | 1860  |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca   | gatatcgcgg | ccgctctaga | caccatgcgc  | 1920  |
| gtgaaggaga | agtaccagca | cctataacac   | tggggctggc | gctggggcac | catgctgctg  | 1980  |
| ggcatgctga | tgatctgcag | caccaccaaa   | aagctgtggg | tgaccgtgta | ctacqqcqtq  | 2040  |
| ccatataa   | aggagggag  | caccaccctg   | ctctgcgcca | gcgacgccaa | gacctacgac  | 2100  |
| accasaatac | acascototo | aaccacccac   | gcctgcgtgc | ccaccgaccc | caacccccag  | 2160  |
| accyaggige | taataaacat | ggccaccaac   | ttcgacatgt | ddaadaacda | catogtogag  | 2220  |
| gaggtggtgt | aggagatast | caccetataa   | gaccagagcc | taaaaccta  | catgatagata | 2280  |
| cagacycacy | aggacaccac | cageeegegg   | gaccagagee | acqacaccaa | caccaacac   | 2340  |
| accecectge | gegegageee | gaagegeace   | gacctgaaga | actacaactt | caecatcage  | 2400  |
| agcagcggcc | gcatgatcat | ggagaagggc   | gagatcaaga | tabagagaga | gangatoato  | 2460  |
| accagcatcc | geggeaaggt | gcagaaggag   | tacgccttct | tetacaaget | ggacatcatc  | 2400  |
| cccatcgaca | acgacaccac | cagctacage   | ctgaccagct | gcaacaccag | egtgateaee  | 2520  |
| caggcctgcc | ccaaggtgag | cttcgagccc   | atccccaacc | actactgcgc | ccccaccaac  | 2580  |
| ttcgccatcc | tgaagtgcaa | ggacaagaag   | ttcaacggca | agggcccctg | caccaacgtg  | 2640  |
| agcaccgtgc | agtgcaccca | cggcatccgc   | cccgtggtga | gcacccagct | gctgctgaac  | 2700  |
| ggtagcctgg | ccgaggagga | ggtggtgatc   | cgcagcgcta | acttcgccga | caacgccaag  | 2760  |
| gtgatcatcg | tgcagctgaa | cgagagcgtg   | gagatcaact | gcacccgccc | caacaacaac  | 2820  |
| acccgcaaga | gcatccacat | cggccccggc   | cgcgccttct | acaccaccgg | cgagatcatc  | 2880  |
| ggcgacatcc | gccaggccca | ctgcaacctg   | ageegegeea | agtggaacga | caccctgaac  | 2940  |
| aagatcgtga | tcaagctgcg | cgagcagttc   | ggcaacaaga | ccatcgtgtt | caagcacagc  | 3000  |
| agcggcggcg | accccgagat | cgtgacccac   | agcttcaact | gcggcggcga | gttcttctac  | 3060  |
| tgcaacagca | cccagctgtt | caacagcacc   | tggttcaaca | gcacctggag | caccgagggc  | 3120  |
| agcaacaaca | ccgagggcag | cgacaccatc   | accctgccct | gccgcatcgc | tagcagcaac  | 3180  |
| aacqaqaqcq | agatetteeg | tccgggcggc   | ggcgacatgc | gcgacaactg | gcgcagcgag  | 3240  |
| ctqtacaaqt | acaaqqtqqt | gaagatcgag   | cccctgggcg | tggcccccac | caaggccaag  | 3300  |
| cttaccatcc | aggcccgcca | gctgctgagc   | ggcatcgtgc | agcagcagaa | caacctgctg  | 3360  |
| cacaccatca | aggcccagca | gcacctgctg   | cagctgaccg | tatagaacat | caagcagctg  | 3420  |
| caddcccdca | ccctaaccat | ggagcgctac   | ctgaaggacc | agcagcagct | cgagcagatc  | 3480  |
| tagaaccaca | ccacctogs  | gaagtagaac   | cgcgagatca | acaactacac | cagochgato  | 3540  |
| cacacccaca | trazanzaz  | - DECECET    | cacgagaaga | acdadcadda | actactaged  | 3600  |
| atagagagag | ccyayyayay | atageactag   | ttcaacatca | ccaactooct | ataatacata  | 3660  |
| aggacaagt  | tastastast | . grggaacrgg | gtgggcctgc | gcatcgtgtt | coccetacta  | 3720  |
| aagutguta  | ceatgategt | tactatacat   | tetaetteee | accetetet  | tatttaccc   | 3780  |
| agcacccgag | gatecagate | . cgccgcgccc | tctagttgcc | statestt   | chastooch   | 3840  |
| cccccgcgc  | cttccttgac | CCCGGaagge   | gccactccca | digidalite | CLaacaaadC  | 2040  |

| gaggaaattg | catcgcattg | tctgagtagg   | tgtcattcta  | ttctgggggg | tggggtgggg | 3900 |
|------------|------------|--------------|-------------|------------|------------|------|
| caggacagca | agggggagga | . ttgggaagac | aatagcaggc  | atgctgggga | tgcqqtqqq  | 3960 |
| tctatgggta | cccaggtgct | . gaagaattga | cccggttcct  | cctgggccag | aaaqaaqcaq | 4020 |
| gcacatcccc | ttctctgtga | . cacaccctgt | ccacgcccct  | ggttcttagt | tccagcccca | 4080 |
| ctcataggac | actcatagct | caggagggct   | ccgccttcaa  | teccaceege | taaagtactt | 4140 |
| ggagcggtct | ctccctccct | catcagccca   | ccaaaccaaa  | cctagcctcc | aagagtggga | 4200 |
| agaaattaaa | gcaagatagg | ctattaagtg   | cagagggaga  | gaaaatgcct | ccaacatgtg | 4260 |
| aggaagtaat | gagagaaatc | atagaatttt   | aaggccatca  | tggccttaat | cttccgcttc | 4320 |
| ctcgctcact | gactcgctgc | gctcggtcgt   | teggetgegg  | cgagcggtat | cageteacte | 4380 |
| aaaggcggta | atacggttat | ccacagaatc   | aggggataac  | gcaggaaaga | acatgtgage | 4440 |
| aaaaggccag | caaaaggcca | ggaaccgtaa   | aaaggccgcg  | ttgctggcgt | ttttccatag | 4500 |
| gctccgcccc | cctgacgage | atcacaaaaa   | tcgacgctca  | agtcagaggt | ggcgaaaccc | 4560 |
| gacaggacta | taaagatacc | aggcgtttcc   | ccctggaagc  | tecetegtge | actctcctat | 4620 |
| tccgaccctg | ccgcttaccg | gatacctgtc   | cgcctttete  | ccttcgggaa | gcgtggcgct | 4680 |
| ttctcatagc | tcacgctgta | ggtatctcag   | ttcggtgtag  | gtcgttcgct | ccaagetqqq | 4740 |
| ctgtgtgcac | gaaccccccg | ttcagcccga   | ccgctgcgcc  | ttatccggta | actategtet | 4800 |
| tgagtccaac | ccggtaagac | acgacttatc   | gccactggca  | gcagccactg | qtaacaqqat | 4860 |
| tagcagagcg | aggtatgtag | gcggtgctac   | agagttcttg  | aagtggtggc | ctaactacgg | 4920 |
| ctacactaga | agaacagtat | ttggtatctg   | cgctctgctg  | aagccagtta | ccttcggaaa | 4980 |
| aagagttggt | agctcttgat | ccggcaaaca   | aaccaccgct  | ggtagcggtg | gtttttttgt | 5040 |
| ttgcaagcag | cagattacgc | gcagaaaaaa   | aggatctcaa  | gaagatcctt | tgatcttttc | 5100 |
| tacggggtct | gacgctcagt | ggaacgaaaa   | ctcacgttaa  | gggattttgg | tcatgagatt | 5160 |
| atcaaaaagg | atcttcacct | agatcctttt   | aaattaaaaa  | tgaagtttta | aatcaatcta | 5220 |
| aagtatatat | gagtaaactt | ggtctgacag   | ttaccaatgc  | ttaatcagtg | aggcacctat | 5280 |
| ctcagcgatc | tgtctatttc | gttcatccat   | agttgcctga  | ctcggggggg | gggggcgctg | 5340 |
| aggtctgcct | cgtgaagaag | gtgttgctga   | ctcataccag  | gcctgaatcg | ccccatcatc | 5400 |
| cagccagaaa | gtgagggagc | cacggttgat   | gagagctttg  | ttgtaggtgg | accagttggt | 5460 |
| gattttgaac | ttttgctttg | ccacggaacg   | gtctgcgttg  | tcgggaagat | gcgtgatctq | 5520 |
| atccttcaac | tcagcaaaag | ttcgatttat   | tcaacaaagc  | cgccgtcccg | tcaagtcagc | 5580 |
| gtaatgctct | gccagtgtta | caaccaatta   | accaattctg  | attagaaaaa | ctcatcgage | 5640 |
| atcaaatgaa | actgcaattt | attcatatca   | ggattatcaa  | taccatattt | ttgaaaaagc | 5700 |
| cgtttctgta | atgaaggaga | aaactcaccg   | aggcagttcc  | ataggatggc | aagatcctgg | 5760 |
| tateggtetg | cgattccgac | tegtecaaca   | tcaatacaac  | ctattaattt | cccctcgtca | 5820 |
| aaaataaggt | tatcaagtga | gaaatcacca   | tgagtgacga. | ctgaatccgg | tgagaatggc | 5880 |
| aaaagcttat | gcatttcttt | ccagacttgt   | tcaacaggcc  | agccattacg | ctcgtcatca | 5940 |
| aaatcactcg | catcaaccaa | accettattc   | attcgtgatt  | gcgcctgagc | gagacgaaat | 6000 |
| acgcgatcgc | tgttaaaagg | acaattacaa   | acaggaatcg  | aatgcaaccg | gcgcaggaac | 6060 |
| actgccagcg | catcaacaat | attttcacct   | gaatcaggat  | attcttctaa | tacctggaat | 6120 |
| getgtttee  | cggggatcgc | agtggtgagt   | aaccatgcat  | catcaggagt | acggataaaa | 6180 |
| tgcttgatgg | ccggaagagg | cataaattcc   | gtcagccagt  | ttagtctgac | catctcatct | 6240 |
| gtaacatcat | tggcaacgct | acctttgcca   | tgtttcagaa  | acaactetgg | cgcatcgggc | 6300 |
| treccataca | accgatagat | tgtcgcacct   | gattgcccga  | cattategeg | agcccattta | 6360 |
| tacccatata | aatcagcatc | catgttggaa   | tttaatcgcg  | gcctcgagca | agacgtttcc | 6420 |
| cgttgaatat | ggctcataac | accccttgta   | ttactgttta  | tgtaagcaga | cagttttatt | 6480 |
| grrcargarg | acacattttt | atcttgtgca   | atgtaacatc  | agagattttg | agacacaacg | 6540 |
| Lggctttccc | occcccca.  | ttattgaagc   | atttatcagg  | gttattgtct | catgagcgga | 6600 |
| cacacacttg | aacgtattta | gaaaaacaaa   | caaatagggg  | ttccgcgcac | atttccccga | 6660 |
| adayigeede | ccgacgtcta | agaaaccatt   | actaccatga  | cattaaccta | taaaaatagg |      |
| cycaccacga | ggccctttcg | CC           |             |            |            | 6742 |

<210> 54 <211> 6781

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 54 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcaqattgta ctgaqagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcqqt tttqqcaqta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cqtcaatqqq aqtttqtttt qqcaccaaaa tcaacqqqac tttccaaaat qtcqtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtqcattg gaacqcgqat 1020 teccegtgce aagagtgaeg taagtacege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getataetgt ttttggettg gggeetatae acceegett eettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagaetgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040 cccgtgtgga aggaggccac caccacctg ctctgcgcca gcgacgccaa ggcctacgac 2100 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280 acceccetgt gegtgageet gaagtgeace gacetgaaga acgacaceaa caccaacage 2340 agcagoggoo goatgatcat ggagaagggo gagatcaaga actqoagott caacatcaqo 2400 accagcatoc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460 cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cqtqatcacc 2520 caggootgoo ccaaggtgag ottogagooc atocccaaco actactgogo coccgooggo 2580 ttegecatee tgaagtgeaa ggacaagaag tteaaeggea agggeeeetg caceaacgtg 2640 ageacceptgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700 ggtagcctgg ccgaggagga ggtggtgatc cgcagcgcta acttcgccga caacgccaag 2760 gtgatcatcg tgcagctgaa cgagagcgtg gagatcaact gcacccgccc caacaacaac 2820 accegeaaga geatecacat eggeeeegge egegeettet acaceaeegg egagateate 2880 ggegacatee gecaggeeca etgeaacetg ageegegeea agtqqaacga caccetqaac 2940

|            |            |            |                                       |            | caagcacagc |      |
|------------|------------|------------|---------------------------------------|------------|------------|------|
|            |            |            |                                       |            | gttcttctac |      |
|            |            |            |                                       |            | caccgagggc |      |
|            |            |            |                                       |            | gcagatcatc |      |
|            |            |            |                                       |            | ccagatccgc |      |
| tgcagcagca | acatcaccgg | cctgctgctg | acccgcgacg                            | gcggcaacag | caacaacgag | 3300 |
| agcgagatct | tccgtccggg | cggcggcgct | agcgccaagc                            | ttaccgtcca | ggcccgccag | 3360 |
| ctgctgagcg | gcatcgtgca | gcagcagaac | aacctgctgc                            | gcgccatcga | ggcccagcag | 3420 |
|            |            |            |                                       |            | cctggccgtg |      |
|            |            |            |                                       |            | cacctggatg |      |
|            |            |            |                                       |            | cgaggagagc |      |
|            |            |            |                                       |            | ggccagcctg |      |
|            |            |            |                                       |            | catgatcgtg |      |
|            |            |            |                                       |            | atccagatct |      |
|            |            |            |                                       |            | ttccttgacc |      |
|            |            |            |                                       |            | atcgcattgt |      |
|            |            |            |                                       |            | gggggaggat |      |
|            |            |            |                                       |            | ccaggtgctg |      |
|            |            |            |                                       |            | tctctgtgac |      |
|            |            |            |                                       |            | ctcatagete |      |
|            |            |            |                                       |            | tecetecete |      |
|            |            |            |                                       |            | caagataggc |      |
|            |            |            |                                       |            | agagaaatca |      |
|            |            |            |                                       |            | actcgctgcg |      |
|            |            |            |                                       |            | tacggttatc |      |
|            |            |            |                                       |            | aaaaggccag |      |
|            |            |            |                                       |            | ctgacgagca |      |
|            |            |            |                                       |            | aaagatacca |      |
|            |            |            |                                       |            | cgcttaccgg |      |
|            |            |            |                                       |            | cacgctgtag |      |
|            |            |            |                                       |            | aaccccccgt |      |
|            |            |            |                                       |            | cggtaagaca |      |
|            |            |            |                                       |            | ggtatgtagg |      |
|            |            |            |                                       |            | gaacagtatt |      |
|            |            |            |                                       |            | gctcttgatc |      |
|            |            |            |                                       |            | agattacgcg |      |
|            |            |            |                                       |            | acgctcagtg |      |
|            |            |            |                                       |            | tcttcaccta |      |
|            |            |            |                                       |            | agtaaacttg |      |
|            |            |            |                                       |            | gtctatttcg |      |
|            |            |            |                                       |            | gtgaagaagg |      |
|            |            |            |                                       |            | tgagggagcc |      |
|            |            |            |                                       |            | tttgctttgc |      |
|            |            |            |                                       |            | cagcaaaagt |      |
|            |            |            |                                       |            | ccagtgttac |      |
|            |            |            |                                       |            | ctgcaattta |      |
|            |            |            |                                       |            | tgaaggagaa |      |
|            |            |            |                                       |            | gattccgact |      |
|            |            |            |                                       |            | atcaagtgag |      |
|            |            |            |                                       |            | catttctttc |      |
|            |            |            |                                       |            | atcaaccaaa |      |
|            |            |            |                                       |            | gttaaaagga |      |
|            |            |            |                                       |            | atcaacaata |      |
|            |            |            |                                       |            | ggggatcgca |      |
|            |            |            |                                       |            | cggaagaggc |      |
|            |            |            | J ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ | 2225       | -22~~2~23  |      |

```
ataaattccg tcagccagtt tagtctgace atctcatctg taacatcatt ggcaacgcta 6300 cctttgccat gtttcagaaa caactctgge gcatcggget tcccatacaa tcgatagatt 6360 gtcgcacctg attgcccgac attatcgcga gcccatttat acccatataa atcagcatcc 6420 atgttggaat ttaatcgcgg cctcgagcaa gacgtttccc gttgaatatg gctcataaca 6480 ccccttgtat tactgttat gtaagcagac agttttattg ttcatgatga tatatttta 6540 tcttgtgcaa tgtaacatca gagattttga gacacaacgt ggctttccc cccccccat 6600 tattgaagca tttatcaggg ttattgtcc atgagcggat acatatttga atgtatttag 6660 aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa 6720 gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag gccctttcgt 6780 c
```

```
<210> 55
<211> 6808
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcq qtqtqaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccqccatqtt qacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
cataqtaacq ccaataqqqa ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaceteca 960
taqaaqacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080
tettatqeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tetecggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
aggqqctcat qqtcqctcqq caqctccttg ctcctaacag tggaggccag acttaggcac 1560
aqcacaatqc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
egegecacca gacataatag etgacagact aacagactgt teettteeat gggtetttte 1860
tqcaqtcacc qtcqtcqaca cqtqtgatca gatatcqcqg ccgctctaga atgcgcqtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
```

| tactaataat  | charactact | accoacaacc  | tataaatasa | catatactac | ggcgtgcccg | 2040 |
|-------------|------------|-------------|------------|------------|------------|------|
| tatagaagaa  | adccaccacc | accetactet  | acaccaacca | egegeaceae | tagagagaga | 2040 |
| agatagaaaaa | catatagaaa | accongence  | gegeeagega | egecaaggee | tacgacaccg | 2100 |
| tastastast  | cgcgcgggcc | acceacycet  | gegegeeeae | egaceccaae | ccccaggagg | 2160 |
| tagegerggr  | gaacgtgacc | gagaactteg  | acacgcggaa | gaacgacatg | gtggagcaga | 2220 |
| Lycacgagga  | catcatcage | ctgtgggace  | agagcctgaa | gccctgcgtg | aagctgaccc | 2280 |
| ccctgtgcgt  | gagcctgaag | tgcaccgacg  | ctagcaagaa | ctgcagcttc | aacatcagca | 2340 |
| ccagcatccg  | cggcaaggtg | cagaaggagt  | acgccttctt | ctacaagctg | gacatcatcc | 2400 |
| ccatcgacaa  | cgacaccacc | agctacagcc  | tgaccagctg | caacaccagc | gtgatcaccc | 2460 |
| aggcctgccc  | caaggtgagc | ttcgagccca  | tecceaacca | ctactgcgcc | cccgccggct | 2520 |
| tegecatect  | gaagtgcaag | gacaagaagt  | tcaacggcaa | gggcccctgc | accaacgtga | 2580 |
| gcaccgtgca  | gtgcacccac | ggcatccgcc  | ccgtggtgag | cacccagctg | ctgctgaacg | 2640 |
| gtagcctggc  | cgaggaggag | gtggtgatcc  | gcagcgctaa | cttcgccgac | aacgccaagg | 2700 |
| tgatcatcgt  | gcagctgaac | gagagcgtgg  | agatcaactg | cacccgcccc | aacaacaaca | 2760 |
| cccgcaagag  | catccacatc | ggccccggcc  | gcgccttcta | caccaccggc | gagatcatcg | 2820 |
| gcgacatccg  | ccaggcccac | tgcaacctga  | gccgcgccaa | gtggaacgac | accctgaaca | 2880 |
| agatcgtgat  | caagctgcgc | gagcagttcg  | gcaacaagac | catcgtgttc | aagcacagca | 2940 |
| gcggcggcga  | ccccgagatc | gtgacccaca  | gcttcaactg | cggcggcgag | ttcttctact | 3000 |
|             |            |             |            |            | accgagggca |      |
| gcaacaacac  | cgagggcagc | gacaccatca  | ccctqccctq | ccgcatcaag | cagatcatca | 3120 |
| acatqtqqca  | gaaggtgggc | aaggccatgt  | acqcccccc  | catcagcggc | cagatccgct | 3180 |
| qcaqcaqcaa  | catcaccage | ctactactaa  | cccqcqacqq | coocaacaoc | aacaacgaga | 3240 |
| gcgagatett  | ccatccaaac | ggcggcgaca  | tgcgcgacaa | ctaacacaac | gagctgtaca | 3300 |
| agtacaaggt  | ggtgaagatc | gagcccctgg  | acataacccc | caccaaggc  | aagcttaccg | 3360 |
| tccaggcccg  | ccaactacta | agcagcatca  | tacagcagca | gaacaacctg | ctgcgcgcca | 3420 |
| tegaggeeea  | gcagcacctg | ctacaactaa  | ccatatagaa | catcaacceg | ctgcaggccc | 2400 |
| acaccctaac  | catagaacac | tacctgaagg  | accadcadca | actcaagcag | atctggaacc | 3540 |
| acaccaccta  | catagagege | caccegaagg  | traaraarta | gecegageag | atccacagec | 3540 |
|             |            |             |            |            |            |      |
| agtagagga   | cctatagaac | tagttaggaga | tracraacta | ggagetgetg | gagctggaca | 3660 |
| tastastast  | cctgtggaat | stantanaa   | traccaacty | getgeggtae | atcaagctgt | 3720 |
| gagaatagaa  | cycygycygc | ecggegggee  | rgegeategt | greegeegeg | ctgagcatct | 3780 |
| tagattaatt  | accegeracy | cetteetaget | gccagccacc | tgttgtttgt | ccctcccccg | 3840 |
| bbassbass   | gaccccggaa | ggtgccacte  | ccactgteet | ttcctaataa | aatgaggaaa | 3900 |
| ccgcaccgca  | cogcocgage | aggigicati  | ccattetggg | gggrggggrg | gggcaggaca | 3960 |
| gcaaggggga  | ggattgggaa | gacaatagca  | ggcatgctgg | ggatgcggtg | ggctctatgg | 4020 |
| gracecaggi  | getgaagaat | tgaeceggtt  | ccccctgggc | cagaaagaag | caggcacatc | 4080 |
| cecttetetg  | tgacacaccc | tgtccacgcc  | cctäättett | agttccagcc | ccactcatag | 4140 |
|             |            |             |            |            | cttggagcgg |      |
| rererecere  | cctcatcage | ccaccaaacc  | aaacctagcc | tccaagagtg | ggaagaaatt | 4260 |
| aaagcaagat  | aggctattaa | gtgcagaggg  | agagaaaatg | cctccaacat | gtgaggaagt | 4320 |
| aatgagagaa  | atcatagaat | tttaaggcca  | tgatttaagg | ccatcatggc | cttaatcttc | 4380 |
|             |            |             |            |            | cggtatcagc |      |
|             |            |             |            |            | gaaagaacat |      |
|             |            |             |            |            | tggcgttttt |      |
| ccataggctc  | cgccccctg  | acgagcatca  | caaaaatcga | cgctcaagtc | agaggtggcg | 4620 |
| aaacccgaca  | ggactataaa | gataccaggc  | gtttccccct | ggaagctccc | tegtgegete | 4680 |
| tcctgttccg  | accctgccgc | ttaccggata  | cctgtccgcc | tttctccctt | cgggaagcgt | 4740 |
| ggcgctttct  | catagctcac | gctgtaggta  | tctcagttcg | gtgtaggtcg | ttcgctccaa | 4800 |
| gctgggctgt  | gtgcacgaac | ccccgttca   | gcccgaccgc | tgcgccttat | ccggtaacta | 4860 |
| tcgtcttgag  | tccaacccgg | taagacacga  | cttatcgcca | ctggcagcag | ccactggtaa | 4920 |
| caggattagc  | agagcgaggt | atgtaggcgg  | tgctacagag | ttcttgaagt | ggtggcctaa | 4980 |
| ctacggctac  | actagaagaa | cagtatttgg  | tatctgcgct | ctgctgaagc | cagttacctt | 5040 |
| cggaaaaaga  | gttggtagct | cttgatccgg  | caaacaaacc | accgctggta | gcggtggttt | 5100 |
|             |            |             |            |            | atcctttgat |      |
| cttttctacg  | gggtctgacg | ctcagtggaa  | cgaaaactca | cgttaaggga | ttttggtcat | 5220 |
| gagattatca  | aaaaggatct | tcacctagat  | ccttttaaat | taaaaatgaa | gttttaaatc | 5280 |
|             |            | _           |            | -          | _          |      |

```
aatctaaagt atatatgagt aaacttggte tgacagttac caatgettaa tcagtgagge 5340
acctatetea gegatetyte tatttegtte atccatagtt geetgaeteg ggggggggg 5400
qcqctqaqqt ctqcctcqtg aagaaqqtqt tgctgactca taccaggcct gaatcgcccc 5460
atcatccaqc cagaaagtga gggagccacg gttgatgaga gctttgttgt aggtggacca 5520
gttggtgatt ttgaactttt gctttgccac ggaacggtct gcgttgtcgg gaagatgcgt 5580
gatetgatec theaacteag caaaagtteg atttatteaa caaageegee gteeegteaa 5640
gtcagcgtaa tgctctgcca gtgttacaac caattaacca attctgatta gaaaaactca 5700
tegageatea aatgaaactg caatttatte atateaggat tateaatace atatttttga 5760
aaaagccgtt tctgtaatga aggagaaaac tcaccgaggc agttccatag gatggcaaga 5820
tectqqtate ggtetgegat teegactegt ceaacateaa tacaacetat taattteece 5880
tegteaaaaa taaggttate aagtgagaaa teaccatgag tgacgactga atceggtgag 5940
aatggcaaaa gcttatgcat ttctttccag acttgttcaa caggccagcc attacgctcg 6000
tcatcaaaat cactcgcatc aaccaaaccg ttattcattc gtgattgcgc ctgagcgaga 6060
cgaaatacgc gatcgctgtt aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc 6120
aggaacactg ccagcgcatc aacaatattt tcacctgaat caggatattc ttctaatacc 6180
tggaatgetg ttttcccggg gatcgcagtg gtgagtaacc atgcatcatc aggagtacgg 6240
ataaaatgct tgatggtcgg aagaggcata aattccgtca gccagtttag tctgaccatc 6300
tcatctgtaa catcattggc aacgctacct ttgccatgtt tcagaaacaa ctctggcgca 6360
tegggettee catacaateg atagattgte geacetgatt geeegacatt ategegagee 6420
catttatacc catataaatc agcatccatq ttggaattta atcgcggcct cgagcaagac 6480
gtttcccgtt gaatatggct cataacaccc cttgtattac tgtttatgta agcagacagt 6540
tttattqttc atqatqatat atttttatct tgtgcaatgt aacatcagag attttgagac 6600
acaacqtggc tttcccccc ccccattat tgaagcattt atcagggtta ttgtctcatg 6660
aggggataca tatttqaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt 6720
ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa 6780
                                                                  6808
aataggcqta tcacgaggcc ctttcgtc
```

```
<211> 6769
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattqqcca ttqcatacqt tqtatccata tcataatatg tacatttata ttqqctcatg 300
tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tqcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cqtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
taqaagacac cqqqaccqat ccaqcctccg cggccgggaa cggtgcattg gaacgcggat 1020
```

<210> 56

teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tertatgeat getatactgt ttttggertg gggeetatae acceegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecce 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ecatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 ageacaatge ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegecacca gacataatag etgacagact aacagactgt teethtecat gggteththe 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920 aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040 tgtggaagga ggccaccacc accetgetet gegccagega egccaaggee taegacaceg 2100 aggtgcacaa cgtgtgggcc acccacgcct gcgtgcccac cgaccccaac ccccaggagg 2160 tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220 tqcacqagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280 ccctqtqcqt qaqcctqaaq tqcaccqacc tqaaqaacqa caccaacacc aacaqcaqca 2340 qcqqccqcat gatcatggag aagggcgaga tcaagaactg cagcttcgct agcaccagct 2400 gcaacaccag cgtgatcacc caggcctgcc ccaaggtgag cttcgagccc atccccaacc 2460 actactgege eccegeegge ttegecatee tgaagtgeaa ggacaagaag tteaaeggea 2520 agggeeeetg caecaacgtg agcacegtge agtgeaceea eggeateege eeegtggtga 2580 gcacccaget gctgctgaac ggtagcctgg ccgaggagga ggtggtgatc cgcagcgcta 2640 acttogocga caacgocaag gtgatcatog tgcagotgaa cgagagogtg gagatcaact 2700 gcacccgccc caacaacaac acccgcaaga gcatccacat cggccccggc cgcgccttct 2760 acaccacgg cgagatcatc ggcgacatcc gccaggccca ctgcaacctg agccgcgcca 2820 agtggaacga caccetgaac aagategtga teaagetgeg egageagtte ggeaacaaga 2880 ccategtgtt caageacage ageggeggeg acceegagat egtgacecae agettcaact 2940 geggeggega gttettetae tgeaacagea eccagetgtt caacageace tggtteaaca 3000 geacctggag caccgaggge ageaacaaca ccgagggeag cgacaccatc accctgccct 3060 gccgcatcaa gcagatcatc aacatgtggc agaaggtggg caaggccatg tacgccccc 3120 ccatcagegg ccagatcegc tgcagcagca acatcacegg cctgctgctg accegegacg 3180 gcggcaacag caacaacgag agcgagatet tecgtecggg eggeggegac atgegegaca 3240 actggcgcag cgagctgtac aagtacaagg tggtgaagat cgagcccctg ggcgtggccc 3300 ccaccaaggc caagettacc gtccaggccc gccagetgct gagcggcatc gtgcagcagc 3360 agaacaacct getgegegec ategaggece ageageacet getgeagetg acegtgtggg 3420 geatcaagca getgeaggee egeaceetgg eegtggageg etacetgaag gaceagcage 3480 agetegagea gatetggaac cacaccacet ggatggagtg ggacegegag atcaacaact 3540 acaccagect gatecacage etgategagg agagecagaa ecageaggag aagaaegage 3600 aggagetget ggagetggae aagtgggeea geetgtggaa etggtteaac atcaccaact 3660 ggctgtggta catcaagctg ttcatcatga tegtgggegg cetggtgggc etgegcateg 3720 tgttcgccgt gctgagcatc tgaggatcca gatctgctgt gccttctagt tgccagccat 3780 etgttgtttg cccctcccc gtgccttcct tgaccctgga aggtgccact cccactgtcc 3840 tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg 3900 ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc aggcatgctg 3960 gggatgcggt gggctctatg ggtacccagg tgctgaagaa ttgacccggt tcctcctggg 4020 ccaqaaagaa gcaqqcacat ccccttctct gtgacacacc ctqtccacqc ccctggttct 4080 tagttccagc cccactcata ggacactcat agctcaggag ggctccgcct tcaatcccac 4140 ccgctaaagt acttggagcg gtctctccct ccctcatcag cccaccaaac caaacctagc 4200 ctccaagagt gggaagaaat taaagcaaga taggctatta agtgcagagg gagagaaaat 4260 gcctccaaca tgtgaggaag taatgagaga aatcatagaa ttttaaggcc atgatttaag 4320

```
gecateatgg cettaatett eegetteete geteaetgae tegetgeget eggtegtteg 4380
getgeggega geggtateag eteacteaaa ggeggtaata eggttateea eagaateagg 4440
ggataacqca ggaaaqaaca tqtqaqcaaa aqqccaqcaa aaqqccaqqa accqtaaaaa 4500
ggccgcgttg ctggcgtttt tccatagget ccgccccct gacgagcatc acaaaaatcg 4560
acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc 4620
tggaagetee etegtgeget eteetgttee gaeeetgeeg ettaceggat acetgteege 4680
ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc 4740
ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg 4800
ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc 4860
actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga 4920
gttcttgaag tggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc 4980
totgotgaag coagttacot toggaaaaag agttggtago tottgatoog goaaacaaac 5040
caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 5100
atotoaagaa qatootttga tottttotac qqqgtotqac gotoagtqqa acqaaaacto 5160
acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa 5220
ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta 5280
ccaatgetta atcagtgagg cacctatete agegatetgt etatttegtt catecatagt 5340
tgcctgactc gggggggggg ggcgctgagg tctgcctcgt gaagaaggtg ttgctgactc 5400
ataccaggce tgaatcgccc catcatccag ccagaaagtg agggagccac ggttgatgag 5460
agetttgttg taggtggacc agttggtgat tttgaacttt tgctttgcca cggaacggtc 5520
tgcgttgteg ggaagatgcg tgatctgatc cttcaactca gcaaaagttc gatttattca 5580
acaaagccgc cgtcccgtca agtcagcgta atgctctgcc agtgttacaa ccaattaacc 5640
aattctgatt agaaaaactc atcgagcatc aaatgaaact gcaatttatt catatcagga 5700
ttatcaatac catatttttg aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg 5760
cagttecata ggatggcaag atcetggtat eggtetgega tteegacteg tecaacatea 5820
atacaaccta ttaatttccc ctcgtcaaaa ataaggttat caagtgagaa atcaccatga 5880
qtqacqactq aatccgqtga qaatqqcaaa agcttatqca tttctttcca gacttqttca 5940
acaggecage cattacgete gtcatcaaaa teactegeat caaccaaace gttattcatt 6000
cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt taaaaggaca attacaaaca 6060
ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa 6120
tcaggatatt cttctaatac ctggaatgct gttttcccgg ggatcgcagt ggtgagtaac 6180
catgcatcat caggagtacg gataaaatgc ttgatggtcg gaagaggcat aaattccgtc 6240
agccagttta gtctgaccat ctcatctgta acatcattgg caacgctacc tttgccatgt 6300
ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt cgcacctgat 6360
tgcccgacat tatcgcgagc ccatttatac ccatataaat cagcatccat gttggaattt 6420
aatcgcggcc tcgagcaaga cgtttcccgt tgaatatggc tcataacacc ccttgtatta 6480
ctgtttatgt aagcagacag ttttattgtt catgatgata tatttttatc ttgtgcaatg 6540
taacatcaga gattttgaga cacaacgtgg ctttcccccc ccccccatta ttgaagcatt 6600
tatcaqqqtt attqtctcat qaqcqqatac atatttqaat qtatttaqaa aaataaacaa 6660
ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 6720
atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtc
                                                                  6769
```

```
<210> 57
<211> 6784
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 57
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcg tcagcggtcg 120
```

| ttggcgggtg | teggggetgg  | cttaactatq  | cggcatcaga  | gcagattgta | ctgagagtgc | 180  |
|------------|-------------|-------------|-------------|------------|------------|------|
| accatatgcg | gtgtgaaata  | ccqcacaqat  | gcgtaaggag  | aaaataccgc | atcagattgg | 240  |
| ctattggcca | ttqcatacqt  | totatccata  | tcataatato  | tacatttata | ttggctcatg | 300  |
| tccaacatta | ccqccatqtt  | gacattgatt  | attgactagt  | tattaatagt | aatcaattac | 360  |
| ggggtcatta | gttcatagcc  | catatataga  | attccacatt  | acataactta | cggtaaatgg | 420  |
| cccacctaac | tgaccgccca  | acqacccccq  | cccattgacg  | tcaataatga | cgtatgttcc | 480  |
| catagtaacg | ccaataggga  | ctttccattg  | acatcaataa  | gragagraft | tacggtaaac | 540  |
| tacccactta | gcagtacate  | aagtgtatca  | tatoccaage  | acqcccccta | ttgacgtcaa | 600  |
| tgacggtaaa | taacccacct  | ggcattatgc  | ccagtacato  | accttatoco | actttcctac | 660  |
| ttggcagtac | atctacgtat  | tagtcatcgc  | tattaccato  | atastacaat | tttggcagta | 720  |
| catcaatogo | catagataga  | gatttgactc  | acqqqqattt  | ccaactctcc | accccattga | 720  |
| cotcaatooo | agtttgttt   | ggcacgaeae  | tcaaccccac  | tttccaaaat | gtcgtaacaa | 940  |
| ctccacccca | ttgacgcaaa  | tagacagtag  | acatatacaa  | tagaaaatat | atataagcag | 940  |
| agetegttta | gtgaaccgtc  | agatogootg  | gagagaacgg  | ccacactatt | ttgacctcca | 900  |
| tagaagacac | cadascast   | ccadcctcca  | caaccacas   | coatgastta | gaacgcggat | 1020 |
| tecceatace | aagagtgacg  | taagtaccoc  | chahacachc  | tatagggaga | cccctttggc | 1020 |
| tettateeat | actatactat  | ttttaactta  | gggggtatag  | cacaggcaca | coccuttgge | 1080 |
| taggtgatg  | tatagettage | cetatagette | tagattatta  | acceeegett | ccttatgcta | 1140 |
| tattogtosc | gatagtttag  | attactasts  | cyggically  | accattattg | accactccc  | 1200 |
| tattggtgat | gatactttcc  | tatataatt   | cataacatgg  | ctetttgeca | caactatctc | 1260 |
| cactggctat | acgecaatae  | tetgreette  | agagaetgae  | acggactctg | tatttttaca | 1320 |
| ggatggggte | ccatttatta  | cccacaacc   | cacatataca  | acaacgccgt | ccccgtgcc  | 1380 |
| cycaycctcc | actaaacata  | gegegggate  | tccacgcgaa  | tctcgggtac | gtgttccgga | 1440 |
| catgggetet | ccccggtag   | cggcggagct  | tccacatccg  | agccctggtc | ccatgcctcc | 1500 |
| ageggeteat | ggtcgctcgg  | cagctccttg  | ctcctaacag  | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac  | cagtgtgccg  | cacaaggccg  | tggcggtagg | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg  | ggctcgcacg  | gctgacgcag  | atggaagact | taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag  | ctgagttgtt  | gtattctgat  | aagagtcaga | ggtaactccc | 1740 |
| grtgeggtge | tgttaacggt  | ggagggcagt  | gtagtctgag  | cagtactcgt | tgctgccgcg | 1800 |
| cgcgccacca | gacataatag  | ctgacagact  | aacagactgt  | tcctttccat | gggtcttttc | 1860 |
| tgcagtcacc | gtcgtcgaca  | cgtgtgatca  | gatatcgcgg  | ccgctctaga | atgcgcgtga | 1920 |
| aggagaagta | ccagcacctg  | tggcgctggg  | gctggcgctg. | gggcaccatg | ctgctgggca | 1980 |
| tgctgatgat | ctgcagcgcc  | accgagaagc  | tgtgggtgac  | cgtgtactac | ggcgtgcccg | 2040 |
| tgtggaagga | ggccaccacc  | accetgetet  | gcgccagcga  | cgccaaggcc | tacgacaccg | 2100 |
| aggtgcacaa | cgtgtgggcc  | acccacgcct  | gcgtgcccac  | cgaccccaac | ccccaggagg | 2160 |
| tggtgctggt | gaacgtgacc  | gagaacttcg  | acatgtggaa  | gaacgacatg | gtggagcaga | 2220 |
| tgcacgagga | catcatcage  | ctgtgggacc  | agagcctgaa  | gccctgcgtg | aagctgaccc | 2280 |
| ccctgtgcgt | gagcctgaag  | tgcaccgacc  | tgaagaacga  | caccaacacc | aacagcagca | 2340 |
| gcggccgcat | gatcatggag  | aagggcgaga  | tcaagaactg  | cagcttcaac | atcagcacca | 2400 |
| gcatccgcgg | caaggtgcag  | aaggagtacg  | ccttcttcta  | caagctggac | atcatcccca | 2460 |
| tcgacaacga | caccaccagc  | tacagcctga  | ccagctgcaa  | caccagcgtg | atcacccagg | 2520 |
| cctgccccaa | ggtgagcttc  | gagcccatcc  | ccaaccacta  | ctgcgccccc | gccggcttcg | 2580 |
| ccatcctgaa | gtgcaaggac  | aagaagttca  | acggcaaggg  | cccctgcacc | aacgtgagca | 2640 |
| ccgtgcagtg | cacccacggc  | atccgccccg  | tggtgagcac  | ccagctgctg | ctgaacggta | 2700 |
| gcctggccga | ggaggaggtg  | gtgatccgca  | gcgctaactt  | cgccgacaac | gccaaggtga | 2760 |
| tcatcgtgca | gctgaacgag  | agcgtggaga  | tcaactgcac  | ccgcgctagc | gcccactgca | 2820 |
| acctgagccg | cgccaagtgg  | aacgacaccc  | tgaacaagat  | cgtgatcaag | ctgcgcgagc | 2880 |
| agttcggcaa | caagaccatc  | gtgttcaagc  | acagcagcgg  | cggcgacccc | gagatcgtga | 2940 |
| cccacagett | caactgcggc  | ggcgagttct  | tctactgcaa  | cagcacccag | ctgttcaaca | 3000 |
| gcacctggtt | caacagcacc  | tggagcaccg  | agggcagcaa  | caacaccgag | ggcagcgaca | 3060 |
| ccatcaccct | gccctgccgc  | atcaagcaga  | tcatcaacat  | gtggcagaag | gtgggcaagg | 3120 |
| ccatgtacgc | ccccccatc   | agcggccaga  | tccgctgcag  | cagcaacatc | accggcctgc | 3180 |
| tgctgacccg | cgacggcggc  | aacagcaaca  | acgagagcga  | gatcttccgt | ccgggcggcg | 3240 |
| gcgacatgcg | cgacaactgg  | cgcagcgagc  | tgtacaagta  | caaggtggtg | aagatcgagc | 3300 |
| ccctgggcgt | ggcccccacc  | aaggccaagc  | ttaccgtcca  | ggcccgccag | ctgctgagcg | 3360 |
| gcatcgtgca | gcagcagaac  | aacctgctgc  | gcgccatcga  | ggcccagcag | cacctgctgc | 3420 |
|            |             |             | •           |            |            |      |

|            |            | aagcagctgc |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            | gagcagatct |            |            |            |      |
|            |            | agcctgatcc |            |            |            |      |
|            |            | ctgctggagc |            |            |            |      |
|            |            | tggtacatca |            |            |            |      |
|            |            | gccgtgctga |            |            |            |      |
|            |            | gtttgcccct |            |            |            |      |
|            |            | taataaaatg |            |            |            |      |
|            |            | ggggtggggc |            |            |            |      |
|            |            | gcggtgggct |            |            |            |      |
| ccggttcctc | ctgggccaga | aagaagcagg | cacatcccct | tctctgtgac | acaccctgtc | 4080 |
|            |            | ccagccccac |            |            |            |      |
|            |            | aaagtacttg |            |            |            |      |
|            |            | agagtgggaa |            |            |            |      |
|            |            | caacatgtga |            |            |            |      |
|            |            | catggcctta |            |            |            |      |
|            |            | ggcgagcggt |            |            |            |      |
|            |            | acgcaggaaa |            |            |            |      |
| caggaaccgt | aaaaaggccg | cgttgctggc | gtttttccat | aggctccgcc | cccctgacga | 4560 |
| gcatcacaaa | aatcgacgct | caagtcagag | gtggcgaaac | ccgacaggac | tataaagata | 4620 |
| ccaggcgttt | ccccctggaa | gctccctcgt | gcgctctcct | gttccgaccc | tgccgcttac | 4680 |
| cggatacctg | teegeettte | tcccttcggg | aagcgtggcg | ctttctcata | gctcacgctg | 4740 |
| taggtatctc | agttcggtgt | aggtcgttcg | ctccaagctg | ggctgtgtgc | acgaaccccc | 4800 |
|            |            | ccttatccgg |            |            |            |      |
|            |            | cagcagccac |            |            |            |      |
| aggcggtgct | acagagttct | tgaagtggtg | gcctaactac | ggctacacta | gaagaacagt | 4980 |
|            |            | tgaagccagt |            |            |            |      |
|            |            | ctggtagcgg |            |            |            |      |
|            |            | aagaagatcc |            |            |            |      |
|            |            | aagggatttt |            |            |            |      |
|            |            | aatgaagttt |            |            |            |      |
|            |            | gcttaatcag |            |            |            |      |
|            |            | gactcggggg |            |            |            |      |
|            |            | aggcctgaat |            |            |            |      |
|            |            | tgttgtaggt |            |            |            |      |
|            |            | tgtcgggaag |            |            |            |      |
|            |            | gccgccgtcc |            |            |            |      |
|            |            | tgattagaaa |            |            |            |      |
|            |            | aataccatat |            |            |            |      |
|            |            | ccataggatg |            |            |            |      |
|            |            | acctattaat |            |            |            |      |
|            |            | gactgaatcc |            |            |            |      |
|            |            | ccagccatta |            |            |            |      |
|            |            | ttgcgcctga |            |            |            |      |
|            |            | cgaatgcaac |            |            |            |      |
|            |            | atattcttct |            |            |            |      |
|            |            | atcatcagga |            |            |            |      |
|            |            | gtttagtctg |            |            |            |      |
|            |            | aaacaactct |            |            |            |      |
|            |            | gacattatcg |            |            |            |      |
|            |            | cggcctcgag |            |            |            |      |
|            |            | tatgtaagca |            |            |            |      |
|            |            | tcagagattt |            |            |            |      |
|            |            | gggttattgt |            |            |            |      |
|            |            | ggttccgcgc |            |            |            |      |
| Luguadaca  |            | 222-3-     |            | J          |            |      |

taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 6780 cqtc

```
<210> 58
<211> 6787
<212> DNA
<213> Artificial Sequence
<223> plasmid pVR1012x/s containing HIV genes
<400> 58
tegegegttt eggtgatgae ggtgaaaacc tetgacacat geageteeeg gagaeggtea 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgeg gtgtgaaata cegeacagat gegtaaggag aaaatacege atcagattgg 240
ctattqqcca ttqcatacqt tqtatccata tcataatatg tacatttata ttqqctcatg 300
tccaacatta ccqccatqtt qacattqatt attqactaqt tattaataqt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
eccqcctqqc tqaccqccca acgacccccq cccattgacq tcaataatga cqtatgttcc 480
cataqtaacq ccaataqqqa ctttccattq acqtcaatqq qtqqaqtatt tacqqtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960
tagaagacac cgggaccgat ccageeteeg eggeegggaa eggtgeattg gaaegeggat 1020
teceegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080
tettatgeat getatactgt tittggettg gggeetatae acceegett eettatgeta 1140
taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecce 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
ageggeteat ggtegetegg eageteettg eteetaacag tggaggeeag acttaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcqgtagg gtatgtgtct 1620
gaaaatgage gtggagattg ggetegeacg getgaegeag atggaagaet taaggeageg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tectttecat gggtetttte 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca '1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accetgetet gegccagega egecaaggee tacgacaceg 2100
aggtgcacaa cqtqtggqcc acccacgcct gcgtgcccac cgaccccaac ccccaggagg 2160
tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcaqc ctqtgggacc agagcctgaa gccctqcqtg aaqctgaccc 2280
ccetgtgcgt gagcetgaag tgcacegace tgaagaacga caccaacace aacagcagca 2340
gcggccgcat gatcatggag aagggcgaga tcaagaactg cagcttcaac atcagcacca 2400
gcatccgcgg caaggtgcag aaggagtacg ccttcttcta caagctggac atcatcccca 2460
```

| h          |            |            |            |            |            | 0-00 |
|------------|------------|------------|------------|------------|------------|------|
|            | caccaccage |            |            |            |            |      |
|            | ggtgagcttc |            |            |            |            |      |
|            | gtgcaaggac |            |            |            |            |      |
|            | cacccacggc |            |            |            |            |      |
|            | ggaggaggtg |            |            |            |            |      |
|            | gctgaacgag |            |            |            |            |      |
|            | ccacategge |            |            |            |            |      |
|            | ggcccactgc |            |            |            |            |      |
|            | gctgcgcgag |            |            |            |            |      |
|            | cgagatcgtg |            |            |            |            |      |
|            | cctgccctgc |            |            |            |            |      |
|            | cgccccccc  |            |            |            |            |      |
|            | ccgcgacggc |            |            |            |            |      |
|            | gcgcgacaac |            |            |            |            |      |
|            | cgtggccccc |            |            |            |            |      |
|            | gcagcagcag |            |            |            |            |      |
| tgcagctgac | cgtgtggggc | atcaagcagc | tgcaggcccg | caccctggcc | gtggagcgct | 3480 |
|            | ccagcagcag |            |            |            |            |      |
|            | caacaactac |            |            |            |            |      |
| agcaggagaa | gaacgagcag | gagctgctgg | agctggacaa | gtgggccagc | ctgtggaact | 3660 |
|            | caccaactgg |            |            |            |            |      |
| tggtgggcct | gcgcatcgtg | ttcgccgtgc | tgagcatctg | aggatccaga | tetgetgtge | 3780 |
| cttctagttg | ccagccatct | gttgtttgcc | cctcccccgt | gccttccttg | accctggaag | 3840 |
| gtgccactcc | cactgtcctt | tcctaataaa | atgaggaaat | tgcatcgcat | tgtctgagta | 3900 |
| ggtgtcattc | tattctgggg | ggtggggtgg | ggcaggacag | caagggggag | gattgggaag | 3960 |
|            | gcatgctggg |            |            |            |            |      |
|            | ctcctgggcc |            |            |            |            |      |
|            | ctggttctta |            |            |            |            |      |
|            | aatcccaccc |            |            |            |            |      |
|            | aacctagcct |            |            |            |            |      |
|            | gagaaaatgc |            |            |            |            |      |
|            | gatttaaggc |            |            |            |            |      |
|            | gtcgttcggc |            |            |            |            |      |
|            | gaatcagggg |            |            |            |            |      |
|            | cgtaaaaagg |            |            |            |            |      |
|            | aaaaatcgac |            |            |            |            |      |
|            | tttccccctg |            |            |            |            |      |
|            | ctgtccgcct |            |            |            |            |      |
|            | ctcagttcgg |            |            |            |            |      |
|            | cccgaccgct |            |            |            |            |      |
|            | ttatcgccac |            |            |            |            |      |
|            | gctacagagt |            |            |            |            |      |
|            | atctgcgctc |            |            |            |            |      |
|            | aaacaaacca |            |            |            |            |      |
|            | aaaaaaggat |            |            |            |            |      |
|            | gaaaactcac |            |            |            |            |      |
|            | cttttaaatt |            |            |            |            |      |
|            | gacagttacc |            |            |            |            |      |
|            | tccatagttg |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | gctgactcat |            |            |            |            |      |
|            | ttgatgagag |            |            |            |            |      |
|            | gaacggtctg |            |            |            |            |      |
|            | tttattcaac |            |            |            |            |      |
|            | aattaaccaa | recegactag | dadaactcat | cgagcaccaa | acgaaactgc | 5700 |
| aacccaccca | tatcaggatt | accaacacca | catttttgaa | aaagccgttt | ccgtaatgaa | 5760 |

```
ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt 5820
ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat aaggttatca 5880
agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag cttatgcatt 5940
tetttecaqa ettqtteaac aggecageca ttacgetegt cateaaaate actegeatea 6000
accaaaccqt tattcattcq tqattgcgcc tgagcgagac gaaatacgcg atcgctgtta 6060
aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 6120
acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 6180
atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaatgctt gatggtcgga 6240
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6300
acqctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6360
tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca 6420
gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg aatatggctc 6480
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6540
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct ttccccccc 6600
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6660
atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6720
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6780
```

```
<210> 59
<211> 6691
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggeege teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tqacqqtaaa tqqcccqcct qqcattatqc ccaqtacatg accttatqqq actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatqqq cqtqqataqc qgtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tecceqtqcc aaqaqtqacq taagtacege etatagacte tatagqcaca eccetttggc 1080
tettatgeat getataetgt ttttggettg gggeetatae acceegett cettatgeta 1140
taggtgatqq tataqcttaq cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
qqatqqqtc ccatttatta tttacaaatt cacatataca acaacqccgt cccccqtqcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteogqtag eggeggaget tecacateeg agecetggte ecatgeetee 1500
```

| agcggctcat | ggtcgctcgg | cageteettg | ctcctaacaq | tqqaggccag | acttaggcac | 1560 |
|------------|------------|------------|------------|------------|------------|------|
|            | ccaccaccac |            |            |            |            |      |
|            | gtggagattg |            |            |            |            |      |
|            | atgcaggcag |            |            |            |            |      |
|            | tgttaacggt |            |            |            |            |      |
|            | gacataatag |            |            |            |            |      |
|            | gtcgtcgaca |            |            |            |            |      |
|            | ccagcacctg |            |            |            |            |      |
|            | ctgcagcgcc |            |            |            |            |      |
|            | ggccaccacc |            |            |            |            |      |
|            | cgtgtgggcc |            |            |            |            |      |
|            | gaacgtgacc |            |            |            |            |      |
|            | catcatcage |            |            |            |            |      |
|            | gagcctgaag |            |            |            |            |      |
|            | ccccaaggtg |            |            |            |            |      |
|            | cctgaagtgc |            |            |            |            |      |
|            | gcagtgcacc |            |            |            |            |      |
| accetaccet | geagegeace | gaggtagta  | teccegege  | tagetteese | gagaagaga  | 2520 |
|            | ggccgaggag |            |            |            |            |      |
|            | cgtgcagctg |            |            |            |            |      |
|            | gagcatccac |            |            |            |            |      |
|            | ccgccaggcc |            |            |            |            |      |
|            | gatcaagctg |            |            |            |            |      |
| gcagcggcgg | cgaccccgag | accgcgaccc | acagetteaa | cracaacaac | gagttettet | 2880 |
| actgcaacag | cacccagctg | ttcaacagca | cctggttcaa | cagcacctgg | agcaccgagg | 2940 |
| gcagcaacaa | caccgagggc | agcgacacca | tcaccctgcc | ctgccgcatc | aagcagatca | 3000 |
|            | gcagaaggtg |            |            |            |            |      |
| gctgcagcag | caacatcacc | ggcctgctgc | tgacccgcga | cggcggcaac | agcaacaacg | 3120 |
|            | cttccgtccg |            |            |            |            |      |
|            | ggtggtgaag |            |            |            |            |      |
| ccgtccaggc | ccgccagctg | ctgagcggca | tcgtgcagca | gcagaacaac | ctgctgcgcg | 3300 |
|            | ccagcagcac |            |            |            |            |      |
|            | ggccgtggag |            |            |            |            |      |
| accacaccac | ctggatggag | tgggaccgcg | agatcaacaa | ctacaccagc | ctgatccaca | 3480 |
|            | ggagagccag |            |            |            |            |      |
|            | cagcctgtgg |            |            |            |            |      |
|            | gatcgtgggc |            |            |            |            |      |
|            | cagatctgct |            |            |            |            |      |
| ccgtgccttc | cttgaccctg | gaaggtgcca | ctcccactgt | cctttcctaa | taaaatgagg | 3780 |
| aaattgcatc | gcattgtctg | agtaggtgtc | attctattct | ggggggtggg | gtggggcagg | 3840 |
| acagcaaggg | ggaggattgg | gaagacaata | gcaggcatgc | tggggatgcg | gtgggctcta | 3900 |
|            | ggtgctgaag |            |            |            |            |      |
| atccccttct | ctgtgacaca | ccctgtccac | gcccctggtt | cttagttcca | gccccactca | 4020 |
|            | atagctcagg |            |            |            |            |      |
| cggtctctcc | ctccctcatc | agcccaccaa | accaaaccta | gcctccaaga | gtgggaagaa | 4140 |
| attaaagcaa | gataggctat | taagtgcaga | gggagagaaa | atgcctccaa | catgtgagga | 4200 |
| agtaatgaga | gaaatcatag | aattttaagg | ccatgattta | aggccatcat | ggccttaatc | 4260 |
|            | tcgctcactg |            |            |            |            |      |
|            | aaggcggtaa |            |            |            |            |      |
|            | aaaggccagc |            |            |            |            |      |
|            | ctccgcccc  |            |            |            |            |      |
|            | acaggactat |            |            |            |            |      |
| ctctcctgtt | ccgaccctgc | cgcttaccgg | atacctgtcc | gectttetee | cttcqqqaaq | 4620 |
| cgtggcgctt | tctcatagct | cacgctgtag | gtatctcagt | teggtataga | tegttegete | 4680 |
|            | tgtgtgcacg |            |            |            |            |      |
|            | gagtccaacc |            |            |            |            |      |
| =          |            |            | _          | 333        | J J 5      |      |

```
taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc 4860
taactacggc tacactagaa gaacagtatt tggtatctgc gctctgctga agccagttac 4920
cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg 4980
tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 5040
gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt 5100
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa 5160
atcaatctaa agtatatatq agtaaacttq qtctqacaqt taccaatqct taatcaqtqa 5220
ggcacctate traggggatet gtetatiteg ticaterata gtiggetgar traggggggg 5280
ggggggctga ggtctgcctc gtgaagaagg tgttgctgac tcataccagg cctgaatcgc 5340
cccatcatcc agccagaaag tgagggagcc acggttgatg agagctttgt tgtaggtgga 5400
ccagttggtg attttgaact tttgctttgc cacggaacgg tctgcgttgt cgggaagatg 5460
cgtgatctga tccttcaact cagcaaaagt tcgatttatt caacaaagcc gccgtcccgt 5520
caagtcagcg taatgctctg ccagtgttac aaccaattaa ccaattctga ttagaaaaac 5580
tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatttt 5640
tgaaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca taggatggca 5700
agatectggt aleggtetge gattecgact egtecaacat caatacaace tattaattte 5760
ccctcgtcaa aaataaggtt atcaagtgag aaatcaccat gagtgacgac tgaatccggt 5820
gagaatggca aaagcttatg catttettte cagacttgtt caacaggeca gecattacge 5880
tegteateaa aateaetege ateaaceaaa eegttattea ttegtgattg egeetgageg 5940
agacgaaata cgcgatcgct gttaaaagga caattacaaa caggaatcga atgcaaccgg 6000
cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata ttcttctaat 6060
acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcatc atcaggagta 6120
cggataaaat gettgatggt cggaagagge ataaatteeg teagecagtt tagtetgaee 6180
atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa caactctggc 6240
geateggget teccatacaa tegatagatt gtegeacetg attgeeegac attategega 6300
gcccatttat acccatataa atcagcatcc atgttggaat ttaatcgcgg cctcgagcaa 6360
gacgtttccc gttgaatatg gctcataaca ccccttgtat tactgtttat gtaagcagac 6420
agttttattg ttcatgatga tatattttta tcttgtgcaa tgtaacatca gagattttga 6480
gacacaacgt ggctttcccc ccccccat tattgaagca tttatcaggg ttattgtctc 6540
atgageggat acatatttga atgtatttag aaaaataaac aaataggggt teegegcaca 6600
tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat 6660
aaaaataggc gtatcacgag gccctttcgt c
                                                                  6691
```

```
<210> 60
<211> 6727
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
```

<400> 60
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt accataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgacggtaaa tggcccgct ggcattatgc ccagtacatg accttatggg actttcctac 660

| ttggcagtac | atctacgtat | tagtcatcgc | tattaccatg | gtgatgcggt  | tttggcagta | 720  |
|------------|------------|------------|------------|-------------|------------|------|
| catcaatggg | cgtggatagc | ggtttgactc | acggggattt | ccaagtctcc  | accccattga | 780  |
| cgtcaatggg | agtttgttt  | ggcaccaaaa | tcaacgggac | tttccaaaat  | gtcgtaacaa | 840  |
| ctccgcccca | ttgacgcaaa | tgggcggtag | gcgtgtacgg | tgggaggtct  | atataagcag | 900  |
| agctcgttta | gtgaaccgtc | agatcgcctg | gagacgccat | ccacgctgtt  | ttgacctcca | 960  |
| tagaagacac | cgggaccgat | ccagcctccg | cggccgggaa | cggtgcattg  | gaacgcggat | 1020 |
| tccccgtgcc | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca  | cccctttggc | 1080 |
| tcttatgcat | gctatactgt | ttttggcttg | gggcctatac | acccccgctt  | ccttatgcta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg | tgggttattg | accattattg  | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttqcca  | caactatctc | 1260 |
| tattggctat | atgccaatac | tctgtccttc | agagactgac | acqqactctq  | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacaccat  | ccccatacc  | 1380 |
| cgcagttttt | attaaacata | gcgtgggatc | tccacqcqaa | tctcqqqtac  | atattccaaa | 1440 |
| catgggctct | tctccggtag | caacaaaact | tccacatccg | agccctggtc  | ccatocctcc | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tagaagaccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggcgg | taacaataaa  | atatatatat | 1620 |
| qaaaatqaqc | gtggagattg | aactcacaca | actaacacaa | atggaagact  | taagggagg  | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagaggggg   | caaggcagcg | 1740 |
| attacaatac | tgttaacggt | adadacact  | gtactetgac | cagtactact  | tastassass | 1/40 |
| cacacascas | geedaegge  | ggagggcagc | acceptate  | testates    | cgetgeegeg | 1800 |
| taasatasaa | gacataatag | ccyacagacc | aacagactgt | ceetteeat   | gggtettte  | 1860 |
| aggagaagta | gtcgtcgaca | tagagatea  | gataccgegg | ccgctctaga  | atgcgcgtga | 1920 |
| aggagaagta | ccagcacctg | rggegerggg | getggegetg | gggcaccatg  | ctgctgggca | 1980 |
| tgetgatgat | ctgcagcgcc | accgagaage | tgtgggtgac | cgtgtactac  | ggcgtgcccg | 2040 |
| cgcggaagga | ggccaccacc | accetgetet | gcgccagcga | cgccaaggcc  | tacgacaccg | 2100 |
| aggtgcacaa | cgtgtgggcc | acccacgcct | gcgtgcccac | cgaccccaac  | ccccaggagg | 2160 |
| tggtgctggt | gaacgtgacc | gagaacttcg | acatgtggaa | gaacgacatg  | gtggagcaga | 2220 |
| tgcacgagga | catcatcagc | ctgtgggacc | agagcctgaa | gccctgcgtg  | aagctgaccc | 2280 |
| ccctgtgcgt | gagcctgaag | tgcaccgacg | ctagcaagaa | ctgcagcttc  | aacatcagca | 2340 |
| ccagcatccg | cggcaaggtg | cagaaggagt | acgccttctt | ctacaagctg  | gacatcatcc | 2400 |
| ccatcgacaa | cgacaccacc | agctacagcc | tgaccagctg | caacaccagc  | gtgatcaccc | 2460 |
| aggcctgccc | caaggtgagc | ttcgagccca | tccccaacca | ctactgcgcc  | cccgccggct | 2520 |
| tcgccatcct | gaagtgcaag | gacaagaagt | tcaacggcaa | gggcccctgc  | accaacgtga | 2580 |
| gcaccgtgca | gtgcacccac | ggcatccgcc | ccgtggtgag | cacccagctg  | ctgctgaacg | 2640 |
| gtagcctggc | cgaggaggag | gtggtgatcc | gcagcgctaa | cttcgccgac  | aacgccaagg | 2700 |
| tgatcatcgt | gcagctgaac | gagagcgtgg | agatcaactg | cacccgcgct  | agcgcccact | 2760 |
| gcaacctgag | ccgcgccaag | tggaacgaca | ccctgaacaa | gatcgtgatc  | aagctgcgcg | 2820 |
| agcagttcgg | caacaagacc | atcgtgttca | agcacagcag | cggcggcgac  | cccgagatcg | 2880 |
| tgacccacag | cttcaactgc | ggcggcgagt | tcttctactg | caacagcacc  | cagctgttca | 2940 |
| acagcacctg | gttcaacagc | acctggagca | ccgagggcag | caacaacacc  | gagggcagcg | 3000 |
| acaccatcac | cctgccctgc | cgcatcaagc | agatcatcaa | catgtggcag  | aaggtgggca | 3060 |
| aggccatgta | cgccccccc  | atcagcggcc | agatccgctg | cagcagcaac  | atcaccggcc | 3120 |
| tgctgctgac | ccgcgacggc | ggcaacagca | acaacgagag | cgagatette  | cgtccgggcg | 3180 |
| gcggcgacat | gcgcgacaac | tggcgcagcg | agctgtacaa | gtacaaggtg  | gtgaagatcg | 3240 |
| agcccctggg | cgtggccccc | accaaggcca | agcttaccgt | ccaggcccgc  | cagctgctga | 3300 |
| gcggcatcgt | gcagcagcag | aacaacctgc | tgcgcgccat | cqaqqcccaq  | cagcacctgc | 3360 |
| tgcagctgac | cgtgtggggc | atcaagcagc | tgcaggcccq | caccctqqcc  | gtggagggct | 3420 |
| acctgaagga | ccagcagcag | ctcgagcaga | tctqqaacca | caccacctqq  | atqqaqtqqq | 3480 |
| accgcgagat | caacaactac | accagcctga | tccacagcct | gatcgaggag  | agccagaacc | 3540 |
| agcaggagaa | gaacgagcag | gagetgetge | agctggacaa | gtgggccagc  | ctgtggaach | 3600 |
| ggttcaacat | caccaactgg | ctgtggtaca | tcaagetgtt | catcatgatc  | atagacaaca | 3660 |
| tggtgggcct | gcgcatcgtg | ttcgccqtqc | tgagcatcto | aggatecaga  | tetachatan | 3720 |
| cttctagttg | ccagccatct | gttgtttacc | cctcccccat | gccttcctta  | accetagaag | 3780 |
| gtgccactcc | cactgtcctt | tcctaataaa | atgaggaaat | tacateacat  | tatctaaata | 3840 |
| ggtgtcattc | tattctgggg | ggtgggataa | ggcaggacag | caagagagag  | gattaggaag | 3900 |
| acaataqcaq | gcatgctggg | gatgcggtag | gctctataga | tacccagata  | ctgaagaath | 3960 |
|            |            | 2 2 22-25  | J          |             |            | 2200 |

```
gacceggtte etectgggee agaaagaage aggeacatee eettetetgt gacacaceet 4020
gtccacgccc ctggttctta gttccagccc cactcatagg acactcatag ctcaggaggg 4080
ctccgccttc aatcccaccc gctaaagtac ttggagcggt ctctccctcc ctcatcagcc 4140
caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata ggctattaag 4200
tgcagaggga gagaaaatgc ctccaacatg tgaggaagta atqagagaaa tcatagaatt 4260
ttaaggccat gatttaaggc catcatggcc ttaatcttcc gcttcctcqc tcactqactc 4320
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 4380
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaa gccagcaaaa 4440
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctqa 4500
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 4560
ataccaggeg threccetg gaageteect egtgegetet cetgtteega eeetgeeget 4620
taceggatac etgteegeet ttetecette gggaagegtg gegetttete atageteaeg 4680
ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacqaacc 4740
ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 4800
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 4860
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4920
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttqqtaqctc 4980
ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca aqcaqcaqat 5040
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctqacgc 5100
tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaqqatctt 5160
cacctagate ettttaaatt aaaaatqaaq ttttaaatca atetaaaqta tatatqaqta 5220
aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcaq cqatctqtct 5280
atttcgttca tccatagttg cctgactcgg ggggggggg cgctgaggtc tgcctcgtga 5340
agaaggtgtt gctgactcat accaggcctg aatcgcccca tcatccagcc agaaagtgag 5400
ggagccaegg ttgatgagag ctttgttgta ggtggaccag ttggtgattt tgaacttttq 5460
ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatcct tcaactcagc 5520
aaaagttcga tttattcaac aaagccgccg tcccgtcaag tcagcgtaat gctctgccag 5580
tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5640
aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa 5700
ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt 5760
cegactegte caacateaat acaacetatt aattteeeet egteaaaaat aaggttatea 5820
agtgagaaat caccatgagt gacgactgaa tooggtgaga atggcaaaag cttatgcatt 5880
tetttecaga ettgtteaac aggecageca ttacgetegt cateaaaate actegeatea 5940
accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 6000
aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 6060
acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 6120
atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaatgctt gatggtcgga 6180
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6240
acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6300
tagattgtog cacctgattg cccgacatta togcgagccc atttataccc atataaatca 6360
geatecatgt tggaatttaa tegeggeete gageaagaeg ttteeegttq aatatqqete 6420
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6480
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct ttccccccc 6540
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6600
atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6660
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6720
tttcqtc
```

<210> 61

<211> 6730

<212> DNA

<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 61 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatqtt qacattqatt attqactaqt tattaatagt aatcaattac 360 qqqqtcatta qttcataqcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 cataqtaacq ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tqcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttqqcaqtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cqtcaatqqq aqtttqtttt qqcaccaaaa tcaacqqqac tttccaaaat qtcqtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcetecg cggccgggaa cggtgcattg gaacgcggat 1020 tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tettatgeat getatactgt ttttggettg gggeetatac acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattqqctat atqccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 qgatqqqqtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cqcaqttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 qcaqaaqaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920 aggagaagta ccaqcacctg tgqcqctggg gctggcgctg gggcaccatg ctgctgggca 1980 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040 tgtggaagga ggccaccacc accetgetet gegccagega egecaaggee tacgacaceg 2100 aggtgcacaa cgtgtgggcc acccacgcct gcgtgcccac cgaccccaac ccccaggagg 2160 tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220 tgcacgagga catcatcago ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280 ccctqtqcqt gagcctgaag tgcaccgacg ctagcaagaa ctgcagcttc aacatcagca 2340 ccaqcatccq cqqcaaqqtq cagaaggagt acgccttctt ctacaagctg gacatcatcc 2400 ccatcgacaa cgacaccacc agctacagcc tgaccagctg caacaccagc gtgatcaccc 2460 aggectgeec caaggtgage ttegageeca tecceaacea etactgegee eeeggeet 2520 tegecatect gaagtgeaag gacaagaagt teaaeggeaa gggeeeetge accaaegtga 2580 gcaccgtgca gtgcacccac ggcatccgcc ccgtggtgag cacccagctg ctgctgaacg 2640 gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcgccgac aacgccaagg 2700 tgatcategt geagetgaae gagagegtgg agatcaaetg caccegeece aacaacaaca 2760 cccgcaagag catccacate ggccccggcc gcgcctteta caccaccggc gagatcatcg 2820 gegacateeg eeaggeeeae tgeaacetga geegegeeaa gtggaacgae accetgaaca 2880 agatogtgat caagotgogo gagoagttog goaacaagao catogtgtto aagoacagoa 2940 qcggcggcga ccccgagatc gtgacccaca gcttcaactg cggcggcgag ttcttctact 3000 gcaacagogo tagootgooo tgoogoatoa agoagatoat caacatgtgg cagaaggtgg 3060

|            | gtacgccccc |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | gacccgcgac |            |            |            |            |      |
|            | catgcgcgac |            |            |            |            |      |
|            | gggcgtggcc |            |            |            |            |      |
|            | cgtgcagcag |            |            |            |            |      |
|            | gaccgtgtgg |            |            |            |            |      |
|            | ggaccagcag |            |            |            |            |      |
|            | gatcaacaac |            |            |            |            |      |
|            | gaagaacgag |            |            |            |            |      |
|            | catcaccaac |            |            |            |            |      |
|            | cctgcgcatc |            |            |            |            |      |
|            | ttgccagcca |            |            |            |            |      |
|            | tcccactgtc |            |            |            |            |      |
|            | ttctattctg |            |            |            |            |      |
|            | caggcatgct |            |            |            |            |      |
|            | ttcctcctgg |            |            |            |            |      |
|            | cccctggttc |            |            |            |            |      |
|            | ttcaatccca |            |            |            |            |      |
| gcccaccaaa | ccaaacctag | cctccaagag | tgggaagaaa | ttaaagcaag | ataggctatt | 4200 |
| aagtgcagag | ggagagaaaa | tgcctccaac | atgtgaggaa | gtaatgagag | aaatcataga | 4260 |
| attttaaggc | catgatttaa | ggccatcatg | gccttaatct | teegetteet | cgctcactga | 4320 |
| ctcgctgcgc | tcggtcgttc | ggctgcggcg | agcggtatca | gctcactcaa | aggcggtaat | 4380 |
|            | acagaatcag |            |            |            |            |      |
| aaaggccagg | aaccgtaaaa | aggccgcgtt | gctggcgttt | ttccataggc | teegeeeeee | 4500 |
| tgacgagcat | cacaaaaatc | gacgctcaag | tcagaggtgg | cgaaacccga | caggactata | 4560 |
| aagataccag | gcgtttcccc | ctggaagctc | cctcgtgcgc | tctcctgttc | cgaccctgcc | 4620 |
| gcttaccgga | tacctgtccg | cctttctccc | ttcgggaagc | gtggcgcttt | ctcatagctc | 4680 |
| acgctgtagg | tatctcagtt | cggtgtaggt | cgttcgctcc | aagctgggct | gtgtgcacga | 4740 |
| accccccgtt | cagcccgacc | gctgcgcctt | atccggtaac | tatcgtcttg | agtccaaccc | 4800 |
| ggtaagacac | gacttatcgc | cactggcagc | agccactggt | aacaggatta | gcagagcgag | 4860 |
| gtatgtaggc | ggtgctacag | agttcttgaa | gtggtggcct | aactacggct | acactagaag | 4920 |
| aacagtattt | ggtatctgcg | ctctgctgaa | gccagttacc | ttcggaaaaa | gagttggtag | 4980 |
|            | ggcaaacaaa |            |            |            |            |      |
| gattacgcgc | agaaaaaaag | gatctcaaga | agatcctttg | atcttttcta | cggggtctga | 5100 |
| cgctcagtgg | aacgaaaact | cacgttaagg | gattttggtc | atgagattat | caaaaaggat | 5160 |
| cttcacctag | atccttttaa | attaaaaatg | aagttttaaa | tcaatctaaa | gtatatatga | 5220 |
| gtaaacttgg | tctgacagtt | accaatgctt | aatcagtgag | gcacctatct | cagcgatctg | 5280 |
| tctatttcgt | tcatccatag | ttgcctgact | cggggggggg | gggcgctgag | gtctgcctcg | 5340 |
| tgaagaaggt | gttgctgact | cataccaggc | ctgaatcgcc | ccatcatcca | gccagaaagt | 5400 |
| gagggagcca | cggttgatga | gagctttgtt | gtaggtggac | cagttggtga | ttttgaactt | 5460 |
| ttgctttgcc | acggaacggt | ctgcgttgtc | gggaagatgc | gtgatctgat | ccttcaactc | 5520 |
| agcaaaagtt | cgatttattc | aacaaagccg | ccgtcccgtc | aagtcagcgt | aatgctctgc | 5580 |
| cagtgttaca | accaattaac | caattctgat | tagaaaaact | catcgagcat | caaatgaaac | 5640 |
| tgcaatttat | tcatatcagg | attatcaata | ccatatttt  | gaaaaagccg | tttctgtaat | 5700 |
| gaaggagaaa | actcaccgag | gcagttccat | aggatggcaa | gatcctggta | tcggtctgcg | 5760 |
| attccgactc | gtccaacatc | aatacaacct | attaatttcc | cctcgtcaaa | aataaggtta | 5820 |
| tcaagtgaga | aatcaccatg | agtgacgact | gaatccggtg | agaatggcaa | aagcttatgc | 5880 |
|            | agacttgttc |            |            |            |            |      |
| tcaaccaaac | cgttattcat | tcgtgattgc | gcctgagcga | gacgaaatac | gcgatcgctg | 6000 |
|            | aattacaaac |            |            |            |            |      |
|            | tttcacctga |            |            |            |            |      |
|            | tggtgagtaa |            |            |            |            |      |
|            | taaattccgt |            |            |            |            |      |
|            | ctttgccatg |            |            |            |            |      |
|            | tegeacetga |            |            |            |            |      |
|            |            |            |            |            |            |      |

```
teageateca tgttggaatt taategegge etegageaag aegttteeeg ttgaatatgg 6420 eteataacae eeettgtatt aetgttatg taageagaea gttttattgt teatgatgat 6480 atattttat ettgtgeaat gtaacateag agattttgag acacaaegtg gettteeee 6540 eeeeeeeatt attgaageat ttateagggt tattgtetea tgageggata eatatttgaa 6600 tgtatttaga aaaataaaca aataggggt eegegeacat tteeeegaaa agtgeeaeet 6660 gaegtetaag aaaccattat tateatgaea ttaacetata aaaataggeg tateaegag 6720 eeetttegte
```

<210> 62
<211> 6688
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVRl012x/s containing HIV genes

<400> 62

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120 ttqqcqqqtq tcqqqqctqq cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 qqqqtcatta qttcatagcc catatatgga qttccqcgtt acataactta cqqtaaatqq 420 cccqcctqqc tgaccgccca acgaccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatea tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acceegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 qqatqqqqtc ccatttatta tttacaaatt cacatataca acaacqccqt cccccqtqcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 aggggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgage gtggagattg ggetegeaeg getgaegeag atggaagaet taaggeageg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegecacea gacataatag etgacagaet aacagaetgt teettteeat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920 aggagaaqta ccaqcacctg tggcqctggq qctqqcqctq qqqcaccatq ctqctqqqca 1980 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040 tgtggaagga ggccaccacc accetgetet gegccagega egccaaqqee taegacaceg 2100 aggtgcacaa cgtgtgggcc acccacgcct gcgtgcccac cgaccccaac ccccaggagg 2160

|   | W U UZ/3294 | 3          |            |            |                          | PC1/U      | 301/23/2 |
|---|-------------|------------|------------|------------|--------------------------|------------|----------|
|   | tggtgctggt  | gaacgtgacc | gagaacttcg | acatgtggaa | gaacgacatg               | gtggagcaga | 2220     |
|   | tgcacgagga  |            |            |            |                          |            |          |
|   | ccctgtgcgt  |            |            |            |                          |            |          |
|   | gcggccgcat  |            |            |            |                          |            |          |
|   |             |            |            |            | cttcgagccc               |            |          |
| • |             |            |            |            | ggacaagaag               |            |          |
|   |             |            |            |            | cggcatccgc               |            |          |
|   | gcacccagct  |            |            |            |                          |            |          |
|   |             |            |            |            | cgagagcgtg               |            |          |
|   |             |            |            |            | gtggaacgac               |            |          |
|   |             |            |            |            | catcgtgttc               |            |          |
|   | geggeggega  |            |            |            |                          |            |          |
|   | gcaacagcac  |            |            |            |                          |            |          |
|   | gcaacaacac  | cgagggcagc | gacaccatca | ccctgccctg | ccgcatcaag               | cagatcatca | 3000     |
|   |             |            |            |            | catcagcggc               |            |          |
|   |             |            |            |            | cggcaacagc               |            |          |
|   |             |            |            |            | ctggcgcagc               |            |          |
|   | agtacaaggt  | ggtgaagatc | gagcccctgg | gcgtggcccc | caccaaggcc               | aagcttaccg | 3240     |
|   | tccaggcccg  | ccagctgctg | agcggcatcg | tgcagcagca | gaacaacctg               | ctgcgcgcca | 3300     |
|   | tcgaggccca  | gcagcacctg | ctgcagctga | ccgtgtgggg | catcaagcag               | ctgcaggccc | 3360     |
|   | gcaccctggc  | cgtggagcgc | tacctgaagg | accagcagca | gctcgagcag               | atctggaacc | 3420     |
|   | acaccacctg  | gatggagtgg | gaccgcgaga | tcaacaacta | caccagectg               | atccacagcc | 3480     |
|   |             |            |            |            | ggagctgctg               |            |          |
|   |             |            |            |            | gctgtggtac               |            |          |
|   | tcatcatgat  | cgtgggcggc | ctggtgggcc | tgcgcatcgt | gttcgccgtg               | ctgagcatct | 3660     |
|   | gaggatccag  | atctgctgtg | ccttctagtt | gccagccatc | tgttgtttgc               | ccctccccg  | 3720     |
|   |             |            |            |            | ttcctaataa               |            |          |
|   |             |            |            |            | gggtggggtg               |            |          |
|   |             |            |            |            | ggatgcggtg               |            |          |
|   |             |            |            |            | cagaaagaag               |            |          |
|   |             |            |            |            | agttccagcc               |            |          |
|   |             |            |            |            | cgctaaagta               |            |          |
|   |             |            |            |            | tccaagagtg               |            |          |
|   |             |            |            |            | cctccaacat               |            |          |
|   |             |            |            |            | ccatcatggc               |            |          |
|   |             |            |            |            | ctgcggcgag               |            |          |
|   |             |            |            |            | gataacgcag               |            |          |
|   |             |            |            |            | gccgcgttgc               |            |          |
|   |             |            |            |            | cgctcaagtc               |            |          |
|   |             |            |            |            | ggaageteee               |            |          |
|   |             |            |            |            | tttctccctt               |            |          |
|   |             |            |            |            | gtgtaggtcg<br>tgcgccttat |            |          |
|   |             |            |            |            | ctggcagcag               |            |          |
|   |             |            |            |            | ttcttgaagt               |            |          |
|   |             |            |            |            | ctgctgaagc               |            |          |
|   |             |            |            |            | accgctggta               |            |          |
|   |             |            |            |            | tctcaagaag               |            |          |
|   |             |            |            |            | cgttaaggga               |            |          |
|   |             |            |            |            | taaaaatgaa               |            |          |
|   |             |            |            |            | caatgcttaa               |            |          |
|   |             |            |            |            | gcctgactcg               |            |          |
|   |             |            |            |            | taccaggcct               |            |          |
|   |             |            |            |            | gctttgttgt               |            |          |
|   |             |            |            |            | gcgttgtcgg               |            |          |
|   | 2003303406  |            | 5000050000 |            | 2-2299                   | 222-2      | 5.100    |

```
gatetgatec tteaacteag caaaagtteg atttatteaa caaageegee gteeegteaa 5520
gtcagcgtaa tgctctgcca gtgttacaac caattaacca attctgatta gaaaaactca 5580
tegageatea aatgaaactg caatttatte atateaggat tateaatace atatttttga 5640
aaaaqccqtt tctqtaatqa aqqaqaaaac tcaccgaggc agttccatag gatggcaaga 5700
tcctggtatc ggtctgcgat tccgactcgt ccaacatcaa tacaacctat taatttcccc 5760
tegteaaaaa taaggttate aagtgagaaa teaceatgag tgacgaetga ateeggtgag 5820
aatggcaaaa gcttatgcat ttctttccag acttgttcaa caggccagcc attacgctcg 5880
teateaaaat caetegeate aaceaaaceg ttatteatte gtgattgege etgagegaga 5940
cgaaatacgc gatcgctgtt aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc 6000
aggaacactg ccagegeate aacaatattt teacetgaat caggatatte ttetaatace 6060
tggaatgctg ttttcccggg gatcgcagtg gtgagtaacc atgcatcatc aggagtacgg 6120
ataaaatgct tgatggtcgg aagaggcata aattccgtca gccagtttag tctgaccatc 6180
tcatctgtaa catcattggc aacgctacct ttgccatgtt tcagaaacaa ctctggcgca 6240
tegggettee catacaateg atagattgte geacetgatt geeegacatt ategegagee 6300
catttatacc catataaatc agcatccatg ttggaattta atcgcggcct cgagcaagac 6360
gtttcccgtt gaatatggct cataacaccc cttgtattac tgtttatgta agcagacagt 6420
tttattgttc atgatgatat atttttatct tgtgcaatgt aacatcagag attttgagac 6480
acaacqtggc tttccccccc ccccattat tgaagcattt atcagggtta ttgtctcatg 6540
aggggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt 6600
ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa 6660
aataggcgta tcacgaggcc ctttcgtc
```

```
<210> 63
<211> 6691
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 63
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tqcccacttq gcaqtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tqacqqtaaa tqqcccqcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae aeggaetetg tatttttaca 1320
```

|            | •          |            |            |            | _          |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | ccccgtgcc  |      |
|            |            |            |            |            | gtgttccgga |      |
|            |            |            |            |            | ccatgcctcc |      |
|            |            |            |            |            | acttaggcac |      |
|            |            |            |            |            | gtatgtgtct |      |
|            |            |            |            |            | taaggcagcg |      |
|            |            |            |            |            | ggtaactccc |      |
|            |            |            |            |            | tgctgccgcg |      |
|            |            |            |            |            | gggtcttttc |      |
|            |            |            |            |            | atgcgcgtga |      |
|            |            |            |            |            | ctgctgggca |      |
| tgctgatgat | ctgcagcgcc | accgagaagc | tgtgggtgac | cgtgtactac | ggcgtgcccg | 2040 |
|            |            |            |            |            | tacgacaccg |      |
| aggtgcacaa | cgtgtgggcc | acccacgcct | gcgtgcccac | cgaccccaac | ccccaggagg | 2160 |
| tggtgctggt | gaacgtgacc | gagaacttcg | acatgtggaa | gaacgacatg | gtggagcaga | 2220 |
| tgcacgagga | catcatcagc | ctgtgggacc | agagcctgaa | gccctgcgtg | aagctgaccc | 2280 |
|            |            |            |            |            | aacagcagca |      |
| gcggccgcat | gatcatggag | aagggcgaga | tcaagaactg | cagcttcgct | agcaccagct | 2400 |
|            |            |            |            |            | atccccaacc |      |
| actactgcgc | ccccgccggc | ttcgccatcc | tgaagtgcaa | ggacaagaag | ttcaacggca | 2520 |
| agggcccctg | caccaacgtg | agcaccgtgc | agtgcaccca | cggcatccgc | cccgtggtga | 2580 |
| gcacccagct | gctgctgaac | ggtagcctgg | ccgaggagga | ggtggtgatc | cgcagcgcta | 2640 |
| acttcgccga | caacgccaag | gtgatcatcg | tgcagctgaa | cgagagcgtg | gagatcaact | 2700 |
| gcacccgccc | caacaacaac | acccgcaaga | gcatccacat | cggccccggc | cgcgccttct | 2760 |
| acaccaccgg | cgagatcatc | ggcgacatcc | gccaggccca | ctgcaacctg | agccgcgcca | 2820 |
| agtggaacga | caccctgaac | aagatcgtga | tcaagctgcg | cgagcagttc | ggcaacaaga | 2880 |
| ccatcgtgtt | caagcacagc | agcggcggcg | accccgagat | cgtgacccac | agcttcaact | 2940 |
| gcggcggcga | gttcttctac | tgcaacagcg | ctagcctgcc | ctgccgcatc | aagcagatca | 3000 |
| tcaacatgtg | gcagaaggtg | ggcaaggcca | tgtacgcccc | ccccatcagc | ggccagatcc | 3060 |
| gctgcagcag | caacatcacc | ggcctgctgc | tgacccgcga | cggcggcaac | agcaacaacg | 3120 |
| agagcgagat | cttccgtccg | ggcggcggcg | acatgcgcga | caactggcgc | agcgagctgt | 3180 |
| acaagtacaa | ggtggtgaag | atcgagcccc | tgggcgtggc | ccccaccaag | gccaagctta | 3240 |
| ccgtccaggc | ccgccagctg | ctgagcggca | tcgtgcagca | gcagaacaac | ctgctgcgcg | 3300 |
|            |            |            |            |            | cagctgcagg |      |
| cccgcaccct | ggccgtggag | cgctacctga | aggaccagca | gcagctcgag | cagatctgga | 3420 |
| accacaccac | ctggatggag | tgggaccgcg | agatcaacaa | ctacaccagc | ctgatccaca | 3480 |
| gcctgatcga | ggagagccag | aaccagcagg | agaagaacga | gcaggagctg | ctggagctgg | 3540 |
|            |            |            |            |            | tacatcaagc |      |
| tgttcatcat | gatcgtgggc | ggcctggtgg | gcctgcgcat | cgtgttcgcc | gtgctgagca | 3660 |
|            |            |            |            |            | tgcccctccc |      |
|            |            |            |            |            | taaaatgagg |      |
|            |            |            |            |            | gtggggcagg |      |
|            |            |            |            |            | gtgggctcta |      |
|            |            |            |            |            | aagcaggcac |      |
| atccccttct | ctgtgacaca | ccctgtccac | gcccctggtt | cttagttcca | gccccactca | 4020 |
| taggacactc | atagctcagg | agggctccgc | cttcaatccc | acccgctaaa | gtacttggag | 4080 |
| cggtctctcc | ctccctcatc | agcccaccaa | accaaaccta | gcctccaaga | gtgggaagaa | 4140 |
|            |            |            |            |            | catgtgagga |      |
|            |            |            |            |            | ggccttaatc |      |
|            |            |            |            |            | gagcggtatc |      |
|            |            |            |            |            | caggaaagaa |      |
| catgtgagca | aaaggccagc | aaaaggccag | gaaccgtaaa | aaggccgcgt | tgctggcgtt | 4440 |
|            |            |            |            |            | gtcagaggtg |      |
|            |            |            |            |            | ccctcgtgcg |      |
| ctctcctgtt | ccgaccctgc | cgcttaccgg | atacctgtcc | gcctttctcc | cttcgggaag | 4620 |
|            |            |            |            |            |            |      |

```
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc 4680
caagetggge tgtgtgcacg aaccecccgt teageccgac cgctgcgcct tatccggtaa 4740
ctategtett gagtecaace eggtaagaca egacttateg ceaetggeag eagceaetgg 4800
taacaggatt agcagagega ggtatgtagg eggtgetaca gagttettga agtggtggee 4860
taactacggc tacactagaa gaacagtatt tggtatctgc gctctgctga agccagttac 4920
cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtaqcqqtqq 4980
ttttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 5040
gatettttet acggggtetg acgeteagtg gaacgaaaac teacgttaag ggattttggt 5100
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaaqttttaa 5160
atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga 5220
ggcacctate teagegatet gtetattteg tteatecata gttgeetgae teggggggg 5280
ggggcgctga ggtctgcctc gtgaagaagg tgttgctgac tcataccagg cctgaatcgc 5340
cccatcatcc agccagaaag tgagggagcc acggttgatg agagctttgt tgtaggtgga 5400
ccagttggtg attttgaact tttgctttgc cacggaacgg tctgcgttgt cgggaagatg 5460
cgtgatctga tccttcaact cagcaaaagt tcgatttatt caacaaagcc gccgtcccgt 5520
caagtcagcg taatgctctg ccagtgttac aaccaattaa ccaattctga ttagaaaaac 5580
tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatttt 5640
tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca taggatggca 5700
agatectggt ateggtetge gatteegact egteeaacat caatacaace tattaattte 5760
ccctcgtcaa aaataaggtt atcaagtgag aaatcaccat gagtgacgac tgaatccggt 5820.
gagaatggca aaagcttatg catttettte cagacttgtt caacaggeca gecattacge 5880
tegteateaa aateactege ateaaceaaa cegttattea ttegtgattg egeetgageg 5940
agacgaaata cgcgatcgct gttaaaagga caattacaaa caggaatcga atgcaaccqq 6000
cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata ttcttctaat 6060
acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcatc atcaqqaqta 6120
cggataaaat gcttgatggt cggaagaggc ataaattccg tcaqccaqtt taqtctqacc 6180
atctcatctg taacatcatt ggcaacgcta cetttgccat gtttcagaaa caactctggc 6240
gcatcgggct tcccatacaa tcgatagatt gtcgcacctg attgcccgac attatcqcqa 6300
geccatttat acceatataa ateageatee atgttggaat ttaategegg cetegageaa 6360
gacgtttccc gttgaatatg gctcataaca ccccttgtat tactgtttat gtaaqcaqac 6420
agttttattg ttcatgatga tatattttta tcttgtgcaa tgtaacatca gagattttga 6480
gacacaacgt ggctttcccc cccccccat tattgaagca tttatcaggg ttattgtctc 6540
atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca 6600
tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat 6660
aaaaataggc gtatcacgag gccctttcgt c
                                                                  6691
```

```
<211> 6706
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 64
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgg tcagcggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gtcatagcc catatatgga gttccgcgtt acataactta ccggtaaatgg 420
cccgcctggc tgaccgcca acgaccccg cccattgacg tcaataatga cgtatgttcc 480
```

<210> 64

|            |            | ctttccattg               |            |            |            |      |
|------------|------------|--------------------------|------------|------------|------------|------|
|            |            | aagtgtatca               |            |            |            |      |
|            |            | ggcattatgc               |            |            |            |      |
| ttggcagtac | atctacgtat | tagtcatcgc               | tattaccatg | gtgatgcggt | tttggcagta | 720  |
|            |            | ggtttgactc               |            |            |            |      |
| cgtcaatggg | agtttgtttt | ggcaccaaaa               | tcaacgggac | tttccaaaat | gtcgtaacaa | 840  |
| ctccgcccca | ttgacgcaaa | tgggcggtag               | gcgtgtacgg | tgggaggtct | atataagcag | 900  |
| agctcgttta | gtgaaccgtc | agatcgcctg               | gagacgccat | ccacgctgtt | ttgacctcca | 960  |
|            |            | ccagcctccg               |            |            |            |      |
|            |            | taagtaccgc               |            |            |            |      |
| tcttatgcat | gctatactgt | ttttggcttg               | gggcctatac | acccccgctt | ccttatgcta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg               | tgggttattg | accattattg | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc               | cataacatgg | ctctttgcca | caactatctc | 1260 |
| tattggctat | atgccaatac | tctgtccttc               | agagactgac | acggactctg | tatttttaca | 1320 |
|            |            | tttacaaatt               |            |            |            |      |
|            |            | gcgtgggatc               |            |            |            |      |
|            |            | cggcggagct               |            |            |            |      |
|            |            | cagctccttg               |            |            |            |      |
|            |            | cagtgtgccg               |            |            |            |      |
|            |            | ggctcgcacg               |            |            |            |      |
|            |            | ctgagttgtt               |            |            |            |      |
| attacaatac | tottaacoot | ggagggcagt               | gtagtctgag | cagtactcgt | tactaccaca | 1800 |
|            |            | ctgacagact               |            |            |            |      |
|            |            | cgtgtgatca               |            |            |            |      |
|            |            | tggcgctggg               |            |            |            |      |
|            |            | accgagaagc               |            |            |            |      |
|            |            | accetgetet               |            |            |            |      |
|            |            | acccacgcct               |            |            |            |      |
|            |            | gagaacttcg               |            |            |            |      |
|            |            | ctgtgggacc               |            |            |            |      |
|            |            | tgcaccgacc               |            |            |            |      |
|            |            | aagggcgaga               |            |            |            |      |
|            |            | aaggagtacg               |            |            |            |      |
|            |            | tacagcctga               |            |            |            |      |
|            |            | gagcccatcc               |            |            |            |      |
|            |            | aagaagttca               |            |            |            |      |
|            |            | atccgccccg               |            |            |            |      |
|            |            | gtgatccgca               |            |            |            |      |
|            |            |                          |            |            |            |      |
|            |            | agegtggaga<br>aaegacaeee |            |            |            |      |
|            |            | gtgttcaagc               |            |            |            |      |
|            |            |                          |            |            |            |      |
|            |            | ggcgagttct<br>atgtggcaga |            |            |            |      |
|            |            |                          |            |            |            |      |
|            |            | agcagcaaca               |            |            |            |      |
|            |            | gagatettee               |            |            |            |      |
|            |            | tacaaggtgg               |            |            |            |      |
|            |            | caggcccgcc               |            |            |            |      |
|            |            | gaggeceage               |            |            |            |      |
|            |            | accetggeeg               |            |            |            |      |
|            |            | accacctgga               |            |            |            |      |
|            |            | atcgaggaga               |            |            |            |      |
|            |            | tgggccagcc               |            |            |            |      |
|            |            | atcatgatcg               |            |            |            |      |
|            |            | ggatccagat               |            |            |            |      |
| ttgtttgccc | ctcccccgtg | ccttccttga               | ccctggaagg | tgccactccc | actgtccttt | 3780 |

| cctaataaaa | tgaggaaatt | gcatcgcatt | gtctgagtag | gtgtcattct | attctggggg | 3840 |
|------------|------------|------------|------------|------------|------------|------|
| gtggggtggg | gcaggacagc | aagggggagg | attgggaaga | caatagcagg | catgctgggg | 3900 |
| atgcggtggg | ctctatgggt | acccaggtgc | tgaagaattg | acceggttee | tcctgggcca | 3960 |
| gaaagaagca | ggcacatccc | cttctctgtg | acacaccctg | tccacgcccc | tggttcttag | 4020 |
| ttccagcccc | actcatagga | cactcatagc | tcaggagggc | tccgccttca | atcccacccg | 4080 |
| ctaaagtact | tggagcggtc | tctccctccc | tcatcagccc | accaaaccaa | acctagcctc | 4140 |
| caagagtggg | aagaaattaa | agcaagatag | gctattaagt | gcagagggag | agaaaatgcc | 4200 |
| tccaacatgt | gaggaagtaa | tgagagaaat | catagaattt | taaggccatg | atttaaggcc | 4260 |
| atcatggcct | taatcttccg | cttcctcgct | cactgactcg | ctgcgctcgg | tcgttcggct | 4320 |
| gcggcgagcg | gtatcagctc | actcaaaggc | ggtaatacgg | ttatccacag | aatcagggga | 4380 |
| taacgcagga | aagaacatgt | gagcaaaagg | ccagcaaaag | gccaggaacc | gtaaaaaggc | 4440 |
| cgcgttgctg | gcgtttttcc | ataggctccg | ccccctgac  | gagcatcaca | aaaatcgacg | 4500 |
| ctcaagtcag | aggtggcgaa | acccgacagg | actataaaga | taccaggcgt | ttccccctgg | 4560 |
| aageteeete | gtgcgctctc | ctgttccgac | cctgccgctt | accggatacc | tgtccgcctt | 4620 |
| tctcccttcg | ggaagcgtgg | cgctttctca | tagctcacgc | tgtaggtatc | tcagttcggt | 4680 |
| gtaggtcgtt | cgctccaagc | tgggctgtgt | gcacgaaccc | cccgttcagc | ccgaccgctg | 4740 |
| caccttatcc | ggtaactatc | gtcttgagtc | caacccggta | agacacgact | tatcgccact | 4800 |
| ggcagcagcc | actggtaaca | ggattagcag | agcgaggtat | gtaggcggtg | ctacagagtt | 4860 |
| cttgaagtgg | tagcctaact | acqqctacac | tagaagaaca | gtatttggta | tetgegetet | 4920 |
| gctgaagcca | gttaccttcg | qaaaaaqagt | tggtagctct | tgatccggca | aacaaaccac | 4980 |
| cactagtage | gatagttttt | ttgtttgcaa | gcagcagatt | acgcgcagaa | aaaaaggatc | 5040 |
| tcaagaagat | cctttgatct | tttctacqqq | gtctgacgct | cagtggaacg | aaaactcacg | 5100 |
| ttaagggatt | ttggtcatga | gattatcaaa | aaggatette | acctagatcc | ttttaaatta | 5160 |
| aaaatgaagt | tttaaatcaa | tctaaaqtat | atatgagtaa | acttggtctg | acagttacca | 5220 |
| atocttaato | agtgaggcac | ctatctcage | gatctgtcta | tttcgttcat | ccatagttgc | 5280 |
| ctgactcggg | aaaaaaaaac | gctgaggtct | gcctcgtgaa | gaaggtgttg | ctgactcata | 5340 |
| ccaggcctga | atcoccccat | catccagcca | gaaagtgagg | gagccacggt | tgatgagagc | 5400 |
| tttattataa | gtggaccagt | tggtgatttt | gaacttttgc | tttgccacgg | aacggtctgc | 5460 |
| attatcaaaa | agatgcgtga | tctgatcctt | caactcagca | aaagttcgat | ttattcaaca | 5520 |
| aagccgcgt  | cccatcaagt | cagcgtaatg | ctctgccagt | gttacaacca | attaaccaat | 5580 |
| tctgattaga | aaaactcatc | gagcatcaaa | tgaaactgca | atttattcat | atcaggatta | 5640 |
| tcaataccat | atttttgaaa | aagccgtttc | totaatgaag | gagaaaactc | accgaggcag | 5700 |
| ttccatagga | tagcaagatc | ctootatcoo | tctgcgattc | cgactcgtcc | aacatcaata | 5760 |
| caacctatta | atttcccctc | otcaaaaata | aggttatcaa | gtgagaaatc | accatgagtg | 5820 |
| acgactgaat | ccaataaaaa | toocaaaagc | ttatgcattt | ctttccagac | ttgttcaaca | 5880 |
| ggccagccat | tacgctcgtc | atcaaaatca | ctcgcatcaa | ccaaaccqtt | attcattcgt | 5940 |
| gattgcgcct | gaggagaga  | aaatacgcga | tcactattaa | aaqqacaatt | acaaacagga | 6000 |
| atcgaatgca | acconcocan | gaacactgcc | agcgcatcaa | caatattttc | acctgaatca | 6060 |
| ggatattett | chaatacctg | gaatgctgtt | ttcccgggga | tcgcagtggt | gagtaaccat | 6120 |
| ggataateaa | gagtacggat | aaaatgettg | atggtcggaa | gaggcataaa | ttccgtcagc | 6180 |
| cantttanto | tgaccatctc | atctgtaaca | tcattggcaa | cactaccttt | gccatgtttc | 6240 |
| agaaacaact | chaacacaba | gagettecca | tacaatcgat | agattgtcgc | acctgattgc | 6300 |
| ccgacattat | cacaaaccca | tttataccca | tataaatcag | catccatqtt | ggaatttaat | 6360 |
| cacaacataa | agcaagacgt | ttcccattaa | atatooctca | taacacccct | tgtattactg | 6420 |
| ttatatasa  | cagacagttt | tattottcat | gatgatatat | ttttatctto | tgcaatgtaa | 6480 |
| catcagagat | tttgagagag | aacqtqqctt | tecceecee  | cccattatto | aagcatttat | 6540 |
| caccagagat | atctcataea | cogatacata | tttgaatgta | tttagaaaaa | taaacaaata | 6600 |
| aggggccacc | gcacatttcc | ccassasata | ccacctgacg | totaagaaac | cattattatc | 6660 |
| atgacattee | cctataaaaa | taggcgtatc | acgagggggg | ttcatc     |            | 6706 |
| argucarcaa | CCCACAAAA  | -475-5-466 |            | 5          | •          |      |

<210> 65

<211> 6649

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 65 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcgggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaqattgq 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatege tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaaqtctcc accccattqa 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataaqcaq 900 agctegttta gtgaaccqte agateqeetq qaqacqccat ccacqetqtt ttqaccteca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtqcattq qaacqcqqat 1020 teccegtgee aagagtgaeg taagtacege etatagaete tataggeaea eecetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acccccqett cettatqeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 taltggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetq tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccqtqcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ecatgeetee 1500 agoggeteat ggtegetegg cageteettg etectaacag tggaggeeag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920 aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040 tgtggaagga ggccaccacc accetgetet gegccagega egccaaggee tacgacaceg 2100 aggtgcacaa cgtgtgggcc acccacgcct gcgtgcccac cgaccccaac ccccaggagg 2160 tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220 tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280 ccctgtgcgt gagcctgaag tgcaccgacg ctagcaagaa ctgcagcttc aacatcagca 2340 ccagcatccg cggcaaggtg cagaaggagt acgccttctt ctacaagctg gacatcatcc 2400 ccategacaa egacaceace agetacagee tgaceagetg caacaceage qtgateacee 2460 aggeotgece caaggtgage thegageeca tecceaacea chachgegee eeegeeggeb 2520 tegecateet gaagtgeaag gacaagaagt teaaeggeaa qqqeecetqe accaaeqtqa 2580 gcaccgtgca gtgcacccac ggcatccgcc ccgtggtgag cacccagctg ctgctgaacg 2640 gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcqccqac aacqccaaqq 2700 tgatcatcgt gcagctgaac gagagcgtgg agatcaactg cacccgcgct agcgcccact 2760 gcaacctgag ccgcgccaag tggaacgaca ccctgaacaa gatcgtgatc aagctgcgcg 2820 agcagttcgg caacaagacc atcgtgttca agcacagcag cggcggcgac cccgagatcg 2880 tgacccacag cttcaactgc ggcggcgagt tcttctactg caacagcgct agcctgccct 2940

| gccgcatcaa | gcagatcatc | aacatgtggc | agaaggtggg  | caaggccatg | tacgccccc  | 3000 |
|------------|------------|------------|-------------|------------|------------|------|
| ccatcagcgg | ccagatccgc | tgcagcagca | acatcaccgg  | cctgctgctg | acccgcgacg | 3060 |
| gcggcaacag | caacaacgag | agcgagatct | tccgtccggg  | cggcggcgac | atgcgcgaca | 3120 |
| actggcgcag | cgagctgtac | aagtacaagg | tggtgaagat  | cgagcccctg | ggcgtggccc | 3180 |
| ccaccaaqqc | caagettace | gtccaggccc | gccagctgct  | gagcggcatc | gtgcagcagc | 3240 |
|            |            | atcgaggccc |             |            |            |      |
| gcatcaagca | gctgcaggcc | cgcaccctgg | ccgtggagcg  | ctacctgaag | gaccagcagc | 3360 |
| agetegagea | gatctggaac | cacaccacct | qqatqqaqtq  | ggaccgcgag | atcaacaact | 3420 |
| acaccagect | gatecacage | ctgatcgagg | agagccagaa  | ccagcaggag | aagaacgagc | 3480 |
| aggagetget | ggagctggac | aagtgggcca | qcctqtqqaa  | ctggttcaac | atcaccaact | 3540 |
| gactataata | catcaagctg | ttcatcatga | tcatagacaa  | cctggtgggc | ctgcgcatcg | 3600 |
| tattcaccat | actgagcatc | tgaggatcca | gatctgctgt  | gccttctagt | tgccagccat | 3660 |
| chattattta | cccctcccc  | gtgccttcct | tgaccctgga  | aggtgccact | cccactqtcc | 3720 |
| tttcctaata | aaatgaggaa | attgcatcgc | attotctgag  | taggtgtcat | tctattctqq | 3780 |
|            |            | agcaaggggg |             |            |            |      |
|            |            | ggtacccagg |             |            |            |      |
| gggacgcgg  | geaggeacat | cccttctct  | ataacacacc  | ctatccacac | ccctggttct | 3960 |
|            |            | ggacactcat |             |            |            |      |
| ccactaaaat | acttonage  | gtctctccct | ccctcatcag  | cccaccaaac | caaacctagc | 4080 |
| ctgccaaagt | acceggageg | taaagcaaga | taggetatta  | agtgcagagg | gagagaaaat | 4140 |
| ccccaagage | totasaasaa | taatgagaga | aatcatacaa  | ttttaaggc  | atgatttaag | 4200 |
| geeceeaca  | cottaatett | ccgcttcctc | actcactgac  | teactagget | caatcattca | 4260 |
| gecateatgg | ccccaatctc | ctcactcaaa | gacaataata  | coattateca | cagacteeg  | 4320 |
| getgeggega | geggeaceag | tgtgagcaaa | aggggggaata | 220002000  | accotasasa | 4380 |
| ggacaacgca | ggaaagaaca | tccataggct | eggccagcaa  | aaggecagga | acegeaaaaa | 4440 |
|            |            |            |             |            |            |      |
| acgeteaagt | cagagguggu | gaaacccgac | aggactataa  | cttaccagg  | acctatacac | 4560 |
| tygaagetee | teresegge  | ctcctgttcc | tastagetas  | cccaccggac | atctgactte | 4620 |
| ctttetecet | cegggaageg | tggcgctttc | tatagacaaa  | coccocatto | accccagecc | 4680 |
|            |            | agetgggetg |             |            |            |      |
|            |            | atcgtcttga |             |            |            |      |
|            |            | acaggattag |             |            |            |      |
|            |            | actacggcta |             |            |            |      |
| cetgetgaag | ccagttacet | toggaaaaag | agttggtagt  | attaggegga | gcaaacaaac | 4920 |
|            |            | tttttgtttg |             |            |            |      |
|            |            | tcttttctac |             |            |            |      |
| acgctaaggg | accttggtca | tgagattatc | aaaaaggacc  | tecaccuaga | ctcccccaa  | 5160 |
| ttaaaaatga | agetttaaat | caatctaaag | tatatatgag  | caaacccggc | ccyacaycca | 2700 |
| ccaatgctta | atcagtgagg | cacctatctc | agegatetgt  | ccaccccgcc | ttaataaata | 5240 |
|            |            | ggcgctgagg |             |            |            |      |
| ataccaggcc | tgaatcgccc | catcatccag | ccagaaagtg  | agggageeae | ggccgacgag | 2340 |
|            |            | agttggtgat |             |            |            |      |
|            |            | tgatctgatc |             |            |            |      |
| acaaagccgc | cgtcccgtca | agtcagcgta | atgetetgee  | agtgttacaa | ccaattaacc | 5520 |
| aattctgatt | agaaaaactc | atcgagcatc | aaatgaaact  | gcaatttatt | catatcagga | 5580 |
| ttatcaatac | catattttg  | aaaaagccgt | ttctgtaatg  | aaggagaaaa | ctcaccgagg | 5640 |
| cagttccata | ggatggcaag | atcctggtat | cggtctgcga  | ttccgactcg | tccaacatca | 5700 |
|            |            | ctcgtcaaaa |             |            |            |      |
| gtgacgactg | aatccggtga | gaatggcaaa | agcttatgca  | cttctttcca | gacttgttca | 5820 |
| acaggccagc | cattacgctc | gtcatcaaaa | ccacccgcat  | caaccaaacc | gctattcatt | 5880 |
| cgtgattgcg | cctgagcgag | acgaaatacg | cgatcgctgt  | taaaaggaca | actacaaaca | 5940 |
|            |            | caggaacact |             |            |            |      |
|            |            | ctggaatgct |             |            |            |      |
|            |            | gataaaatgc |             |            |            |      |
|            |            | ctcatctgta |             |            |            |      |
| ttcagaaaca | actctggcgc | atcgggcttc | ccatacaatc  | gatagattgt | cgcacctgat | 6240 |
|            |            |            |             |            |            |      |

PCT/US01/25721

```
WO 02/32943
tgcccgacat tatcgcgagc ccatttatac ccatataaat cagcatccat gttggaattt 6300
aategeggee tegageaaga egttteeegt tgaatatgge teataacace cettgtatta 6360
ctgtttatgt aagcagacag ttttattgtt catgatgata tatttttatc ttgtgcaatg 6420
taacatcaga gattttgaga cacaacgtgg ctttcccccc cccccatta ttgaagcatt 6480
tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa 6540
ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 6600
atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtc
<210> 66
<211> 6610
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 66
tegegegettt eggegatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagcttgtct gtaageggat geegggagea gacaageeeg teagggegeg teageggqtg 120°
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
```

tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 qqqqtcatta qttcatagcc catatatgga qttccqcqtt acataactta cqqtaaatqq 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaaccgte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtacege etatagaete tataggeaca eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acccegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccqtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeg agecetggte ecatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga qgtaactccc 1740 gttgcggtgc tqttaacqqt qqaqqqcagt gtagtctqaq caqtactcqt tqctqccqcq 1800 egegecacca gacataatag etgacagact aacagactgt teethteeat qqqtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920 aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040 tgtggaagga ggccaccacc accetgetet gcgccagega cgccaaggcc tacgacaccg 2100 aggigacaa cgigtgggcc accaegect gcgtgcccac cgaccccaac ccccaggagg 2160

|            | gaacgtgacc |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | catcatcagc |            |            |            |            |      |
|            | gagcctgaag |            |            |            |            |      |
|            | gatcatggag |            |            |            |            |      |
|            | cgtgatcacc |            |            |            |            |      |
|            | ccccgccggc |            |            |            |            |      |
|            | caccaacgtg |            |            |            |            |      |
|            | gctgctgaac |            |            |            |            |      |
|            | caacgccaag |            |            |            |            |      |
|            | tagcgcccac |            |            |            |            |      |
|            | caagctgcgc |            |            |            |            |      |
|            | ccccgagatc |            |            |            |            |      |
|            | tagcctgccc |            |            |            |            |      |
|            | gtacgccccc |            |            |            |            |      |
|            | gacccgcgac |            |            |            |            |      |
|            | catgcgcgac |            |            |            |            |      |
|            | gggcgtggcc |            |            |            |            |      |
|            | cgtgcagcag |            |            |            |            |      |
|            | gaccgtgtgg |            |            |            |            |      |
|            | ggaccagcag |            |            |            |            |      |
|            | gatcaacaac |            |            |            |            |      |
| accagcagga | gaagaacgag | caggagctgc | tggagctgga | caagtgggcc | agcctgtgga | 3480 |
|            | catcaccaac |            |            |            |            |      |
|            | cctgcgcatc |            |            |            |            |      |
|            | ttgccagcca |            |            |            |            |      |
|            | tcccactgtc |            |            |            |            |      |
|            | ttctattctg |            |            |            |            |      |
|            | caggcatgct |            |            |            |            |      |
|            | ttcctcctgg |            |            |            |            |      |
|            | cccctggttc |            |            |            |            |      |
|            | ttcaatccca |            |            |            |            |      |
|            | ccaaacctag |            |            |            |            |      |
|            | ggagagaaaa |            |            |            |            |      |
|            | catgatttaa |            |            |            |            |      |
|            | teggtegtte |            |            |            |            |      |
|            | acagaatcag |            |            |            |            |      |
|            | aaccgtaaaa |            |            |            |            |      |
|            | cacaaaaatc |            |            |            |            |      |
|            | gcgtttcccc |            |            |            |            |      |
|            | tacctgtccg |            |            |            |            |      |
|            | tatctcagtt |            |            |            |            |      |
|            | cagcccgacc |            |            |            |            |      |
|            | gacttatcgc |            |            |            |            |      |
|            | ggtgctacag |            |            |            |            |      |
|            | ggtatctgcg |            |            |            |            |      |
|            | ggcaaacaaa |            |            |            |            |      |
|            | agaaaaaaag |            |            |            |            |      |
|            | aacgaaaact |            |            |            |            |      |
|            | atccttttaa |            |            |            |            |      |
|            | tctgacagtt |            |            |            |            |      |
|            | tcatccatag |            |            |            |            |      |
|            | gttgctgact |            |            |            |            |      |
|            | cggttgatga |            |            |            |            |      |
|            | acggaacggt |            |            |            |            |      |
|            |            | aacaaagccg | ccatcccatc | aagtcagcgt | aatactctac | 5460 |

```
cagtgttaca accaattaac caattetgat tagaaaaact categagcat caaatgaaac 5520
tgcaatttat tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat 5580
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg 5640
attccqactc gtccaacatc aatacaacct attaatttcc cctcgtcaaa aataaggtta 5700
tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaatggcaa aagcttatgc 5760
atttetttee agaettgtte aacaggeeag ceattaeget egteateaaa ateaetegea 5820
tcaaccaaac cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg 5880
ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 5940
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg 6000
gggatcgcag tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 6060
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6120
gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 6180
cqataqattq tcqcacctga ttqcccgaca ttatcgcgag cccatttata cccatataaa 6240
tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg 6300
ctcataacac cccttgtatt actgtttatg taagcagaca gttttattgt tcatgatgat 6360
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttccccc 6420
ccccccatt attqaaqcat ttatcagggt tattgtctca tgagcggata catatttgaa 6480
tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct 6540
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6600
ccctttcqtc
```

<210> 67 <211> 6610 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes / tegegegettt eggegatgae ggegaaaace teegacacat geageteeeg gagaeggea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tqacqqtaaa tqqcccqcct ggcattatgc ccaqtacatq accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eecetttgge 1080 tottatgcat gotatactgt ttttggcttg gggcctatac acccccgctt cottatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac totgtcottc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380

| caceattttt  | attaaacata | acataaasta | +000000000 | tetegggtag |            | 1440 |
|-------------|------------|------------|------------|------------|------------|------|
| catagagatat | attaaacata | gegegggace | tecacycyaa | cecegggcae | gigiloogga | 1440 |
| acconstant  | tctccggtag | cggcggagct | chactaceeg | agecetggte | ccatgeetee | 1500 |
|             | ggtcgctcgg |            |            |            |            |      |
|             | ccaccaccac |            |            |            |            |      |
| gaaaacgagc  | gtggagattg | ggcccgcacg | gctgacgcag | atggaagact | taaggcagcg | 1680 |
| gcagaagaag  | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
| gttgcggtgc  | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg | 1800 |
| cgcgccacca  | gacataatag | ctgacagact | aacagactgt | tcctttccat | gggtcttttc | 1860 |
| tgcagtcacc  | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | atgcgcgtga | 1920 |
| aggagaagta  | ccagcacctg | tggcgctggg | gctggcgctg | gggcaccatg | ctgctgggca | 1980 |
| tgctgatgat  | ctgcagcgcc | accgagaagc | tgtgggtgac | cgtgtactac | ggcgtgcccg | 2040 |
| tgtggaagga  | ggccaccacc | accctgctct | gcgccagcga | cgccaaggcc | tacgacaccg | 2100 |
|             | cgtgtgggcc |            |            |            |            |      |
| tggtgctggt  | gaacgtgacc | gagaacttcg | acatgtggaa | gaacgacatg | gtggagcaga | 2220 |
| tgcacgagga  | catcatcagc | ctgtgggacc | agagcctgaa | gccctgcgtg | aagctgaccc | 2280 |
|             | gagcctgaag |            |            |            |            |      |
| cccaggcctg  | ccccaaggtg | agcttcgagc | ccatccccaa | ccactactgc | gcccccgccg | 2400 |
| gcttcgccat  | cctgaagtgc | aaggacaaga | agttcaacgg | caagggcccc | tgcaccaacg | 2460 |
|             | gcagtgcacc |            |            |            |            |      |
|             | ggccgaggag |            |            |            |            |      |
| aggtgatcat  | cgtgcagctg | aacgagagcg | tggagatcaa | ctgcacccgc | gctagcgccc | 2640 |
| actgcaacct  | gagccgcgcc | aagtggaacg | acaccctgaa | caagatcgtg | atcaagctgc | 2700 |
| gcgagcagtt  | cggcaacaag | accatcgtgt | tcaagcacag | cagcggcggc | gaccccgaga | 2760 |
| tcgtgaccca  | cagcttcaac | tgcggcggcg | agttcttcta | ctgcaacagc | acccagetgt | 2820 |
|             | ctggttcaac |            |            |            |            |      |
|             | caccctgccc |            |            |            |            |      |
|             | gtacgccccc |            |            |            |            |      |
|             | gacccgcgac |            |            |            |            |      |
|             | catgcgcgac |            |            |            |            |      |
|             | gggcgtggcc |            |            |            |            |      |
| tgagcggcat  | cgtgcagcag | cagaacaacc | tgctgcgcgc | categaggee | cagcagcacc | 3240 |
| tgctgcagct  | gaccgtgtgg | ggcatcaagc | agctgcaggc | ccgcaccctg | gccgtggagc | 3300 |
|             | ggaccagcag |            |            |            |            |      |
| gggaccgcga  | gatcaacaac | tacaccagcc | tgatccacag | cctgatcgag | gagagccaga | 3420 |
| accagcagga  | gaagaacgag | caggagetge | tggagctgga | caagtgggcc | agcctgtgga | 3480 |
|             | catcaccaac |            |            |            |            |      |
|             | cctgcgcatc |            |            |            |            |      |
| tgccttctag  | ttgccagcca | tctgttgttt | gcccctcccc | cgtgccttcc | ttgaccctgg | 3660 |
|             | tcccactgtc |            |            |            |            |      |
| gtaggtgtca  | ttctattctg | gggggtgggg | tggggcagga | cagcaagggg | gaggattggg | 3780 |
| aagacaatag  | caggcatgct | ggggatgcgg | tgggctctat | gggtacccag | gtgctgaaga | 3840 |
|             | ttcctcctgg |            |            |            |            |      |
|             | cccctggttc |            |            |            |            |      |
| gggctccgcc  | ttcaatccca | cccgctaaag | tacttggagc | ggtctctccc | tccctcatca | 4020 |
|             | ccaaacctag |            |            |            |            |      |
|             | ggagagaaaa |            |            |            |            |      |
| attttaaggc  | catgatttaa | ggccatcatg | gccttaatct | teegetteet | cgctcactga | 4200 |
|             | tcggtcgttc |            |            |            |            |      |
| acggttatcc  | acagaatcag | gggataacgc | aggaaagaac | atgtgagcaa | aaggccagca | 4320 |
| aaaggccagg  | aaccgtaaaa | aggccgcgtt | gctggcgttt | ttccataggc | tccgccccc  | 4380 |
| tgacgagcat  | cacaaaaatc | gacgctcaag | tcagaggtgg | cgaaacccga | caggactata | 4440 |
| aagataccag  | gagtttaad  | ctggaagctc | cctcgtgcgc | tctcctgttc | cgaccctgcc | 4500 |
| gcttaccgga  | tacctgtccg | cctttctccc | ttcgggaagc | gtggcgcttt | ctcatagctc | 4560 |
| acgctgtagg  | tatctcagtt | cggtgtaggt | cgttcgctcc | aagctgggct | gtgtgcacga | 4620 |
| accccccgtt  | cagcccgacc | gctgcgcctt | atccggtaac | tatcgtcttg | agtccaaccc | 4680 |
|             |            |            |            |            |            |      |

```
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 4740
qtatqtaqqc qqtqctacag agttcttgaa gtggtggcct aactacggct acactagaag 4800
aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 4860
ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca 4920
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atctttcta cggggtctga 4980
cqctcaqtqq aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 5040
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 5100
gtaaacttgg totgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 5160
tctatttcgt tcatccatag ttgcctgact cggggggggg gggcgctgag gtctgcctcg 5220
tgaagaaggt gttgctgact cataccaggc ctgaatcgcc ccatcatcca gccagaaagt 5280
gagggagcca cggttgatga gagctttgtt gtaggtggac cagttggtga ttttgaactt 5340
ttgctttgcc acggaacggt ctgcgttgtc gggaagatgc gtgatctgat ccttcaactc 5400
agcaaaagtt cgatttatte aacaaageeg eegteeegte aagteagegt aatgetetge 5460
caqtqttaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5520
tgcaatttat tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat 5580
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg 5640
attccgactc gtccaacatc aatacaacct attaatttcc cctcgtcaaa aataaggtta 5700
tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaatggcaa aagcttatgc 5760
atttctttcc agacttgttc aacaggccag ccattacgct cgtcatcaaa atcactcgca 5820
tcaaccaaac cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg 5880
ttaaaaqqac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 5940
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg 6000
gggatcgcag tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 6060
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6120
gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 6180
cgatagattg tegeacetga ttgecegaca ttategegag eccatttata eccatataaa 6240
tcaqcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg 6300
ctcataacac cccttgtatt actgtttatg taagcagaca gttttattgt tcatgatgat 6360
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttccccc 6420
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6480
tqtatttaqa aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct 6540
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6600
                                                                  6610
ccctttcgtc
```

```
<211> 6613
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 68
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagccg tcagggggg tctgagggtg tcggggtg cttaactatg cggcatcaga gcagattgta ctgagagtg 120
ttggcggtt tcggtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
```

<210> 68

| WO 02/3294 | 3          |            |            |            | PCT/US     | 01/257 |
|------------|------------|------------|------------|------------|------------|--------|
| tgacggtaaa | taacccacct | ggcattatgc | ccaqtacatq | accttatggg | actttcctac | 660    |
|            |            |            |            |            | tttggcagta |        |
|            |            |            |            |            | accccattga |        |
|            |            |            |            |            | gtcgtaacaa |        |
| ctccacccca | ttgacgcaaa | tagacaataa | gcgtgtacgg | tgggaggtct | atataagcag | 900    |
| agctcgttta | gtgaaccgtc | agategeetg | qaqacqccat | ccacqctgtt | ttgacctcca | 960    |
|            |            |            |            |            | gaacgcggat |        |
| tececatace | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca | cccctttggc | 1080   |
| tcttatqcat | gctatactgt | ttttggcttg | gggcctatac | acccccgctt | ccttatgcta | 1140   |
| taggtgatgg | tatagettag | cctataggtg | tgggttattg | accattattg | accactcccc | 1200   |
| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260   |
|            |            |            |            |            | tatttttaca |        |
| qqatqqqqtc | ccatttatta | tttacaaatt | cacatataca | acaacgccgt | ccccgtgcc  | 1380   |
|            |            |            |            |            | gtgttccgga |        |
|            |            |            |            |            | ccatgcctcc |        |
|            |            |            |            |            | acttaggcac |        |
|            |            |            |            |            | gtatgtgtct |        |
|            |            |            |            |            | taaggcagcg |        |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740   |
|            |            |            |            |            | tgctgccgcg |        |
|            |            |            |            |            | gggtcttttc |        |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | atgcgcgtga | 1920   |
| aggagaagta | ccaqcacctq | tggcgctggg | getggegetg | gggcaccatg | ctgctgggca | 1980   |
| tgctgatgat | ctgcagcgcc | accgagaagc | tgtgggtgac | cgtgtactac | ggcgtgcccg | 2040   |
|            |            |            |            |            | tacgacaccg |        |
|            |            |            |            |            | ccccaggagg |        |
|            |            |            |            |            | gtggagcaga |        |
|            |            |            |            |            | aagctgaccc |        |
|            |            |            |            |            | agcgtgatca |        |
|            |            |            |            |            | gccccgccg  |        |
|            |            |            |            |            | tgcaccaacg |        |
|            |            |            |            |            | ctgctgctga |        |
|            |            |            |            |            | gacaacgcca |        |
|            |            |            |            |            | cccaacaaca |        |
|            |            |            |            |            | ggcgagatca |        |
|            |            |            |            |            | gacaccctga |        |
|            |            |            |            |            | ttcaagcaca |        |
|            |            |            |            |            | gagttcttct |        |
|            |            |            |            |            | tggcagaagg |        |
|            |            |            |            |            | agcaacatca |        |
|            |            |            |            |            | atcttccgtc |        |
|            |            |            |            |            | aaggtggtga |        |
|            |            |            |            |            | gcccgccagc |        |
|            |            |            |            |            | gcccagcagc |        |
|            |            |            |            |            | ctggccgtgg |        |
|            |            |            |            |            | acctggatgg |        |
|            |            |            |            |            | gaggagagcc |        |
|            |            |            |            |            | gccagcctgt |        |
|            |            |            |            |            | atgatcgtgg |        |
|            |            |            |            |            | tccagatctg |        |
|            |            |            |            |            | tccttgaccc |        |
|            |            |            |            |            | tcgcattgtc |        |
|            |            |            |            |            | ggggaggatt |        |
|            |            |            |            |            | caggtgctga |        |
| agaattgacc | cggttcctcc | tgggccagaa | agaagcaggc | acateceett | ctctgtgaca | 3900   |
|            |            |            |            |            |            |        |

```
caccetgice acgeeeetgg ticttagite cageeeeact cataggacae teatagetea 3960
ggagggetee geetteaate ceaecegeta aagtaettgg ageggtetet ceeteeetea 4020
tcageceace aaaceaaace tageeteeaa gagtgggaag aaattaaage aagatagget 4080
attaagtgca gagggagaga aaatgcctcc aacatgtgag gaagtaatga gagaaatcat 4140
agaattttaa ggccatgatt taaggccatc atggccttaa tcttccgctt cctcgctcac 4200
tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt 4260
aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca 4320
gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc 4380
ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact 4440
ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct 4500
qccqcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag 4560
ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca 4620
cgaacccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa 4680
cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc 4740
qaqqtatqta qqcqqtqcta cagaqttctt gaagtggtgg cctaactacg gctacactag 4800
aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg 4860
tagetettga teeggeaaac aaaceaeege tggtageggt ggtttttttg tttgcaagea 4920
gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc 4980
tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag 5040
qatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata 5100
tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat 5160
ctqtctattt cqttcatcca tagttgcctg actcgggggg ggggggcgct gaggtctgcc 5220
teqtgaaqaa ggtgttgetg acteatacca ggcctgaatc gccccatcat ccagccagaa 5280
aqtqaqqqaq ccacqqttga tgagagcttt gttgtaggtg gaccagttgg tgattttgaa 5340
cttttgcttt gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct gatccttcaa 5400
ctcagcaaaa gttcgattta ttcaacaaag ccgccgtccc gtcaagtcag cgtaatgctc 5460
tgccagtgtt acaaccaatt aaccaattct gattagaaaa actcatcgag catcaaatga 5520
aactgcaatt tattcatatc aggattatca ataccatatt tttgaaaaag ccgtttctgt 5580
aatqaaqqaq aaaactcacc gaggcagttc cataggatgg caagatcctg gtatcggtct 5640
gcgattccga ctcgtccaac atcaatacaa cctattaatt tcccctcgtc aaaaataagg 5700
ttatcaagtg agaaatcacc atgagtgacg actgaatccg gtgagaatgg caaaagctta 5760
tgcatttctt tccagacttg ttcaacaggc cagccattac gctcgtcatc aaaatcactc 5820
qcatcaacca aaccqttatt cattcqtqat tqcgcctgag cgagacgaaa tacgcgatcg 5880
ctqttaaaaq gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc 5940
qcatcaacaa tattttcacc tqaatcagga tattcttcta atacctggaa tgctgttttc 6000
ccggggatcg cagtggtgag taaccatgca tcatcaggag tacggataaa atgcttgatg 6060
gtcggaagag gcataaattc cgtcagccag tttagtctga ccatctcatc tgtaacatca 6120
ttggcaacgc tacctttgcc atgtttcaga aacaactctg gcgcatcggg cttcccatac 6180
aatcqataga ttgtcgcacc tgattgcccg acattatcgc gagcccattt atacccatat 6240
aaatcagcat ccatgttgga atttaatcgc ggcctcgagc aagacgtttc ccgttgaata 6300
tggctcataa caccccttgt attactgttt atgtaagcag acagttttat tgttcatgat 6360
gatatatttt tatcttgtgc aatgtaacat cagagatttt gagacacaac gtggctttcc 6420
ccccccccc attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt 6480
qaatqtattt agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca 6540
cetgaegtet aagaaaceat tattateatg acattaacet ataaaaatag gegtateaeg 6600
                                                                  6613
aggecettte gte
```

<210> 69

<211> 6532

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

```
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtct gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccqccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960
taqaaqacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tettatgeat getatactgt ttttggettg gggcctatac acccccgett cettatgeta 1140
taggtgatgq tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattqqctat atqccaatac tctqtccttc agagactgac acggactctg tatttttaca 1320
qqatqqqqtc ccatttatta tttacaaatt cacatataca acaacqccqt cccccqtqcc 1380
cqcaqttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg ageeetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560
agcacaatgc ccaccacca cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
qaaaatqaqc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagegec acegagaage tgtgggtgae egtgtactae ggegtgeeeg 2040
tgtggaagga ggccaccacc accetgetet gcgccagega cgccaaggee tacgacaccg 2100
aggtgcacaa cgtgtgggcc acccacgcct gcgtgcccac cgaccccaac ccccaggagg 2160
tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcage etgtgggace agageetgaa geeetgegtg aagetgacee 2280
ccctgtgcgt gagcctgaag tgcaccgacg ctagcaccag ctgcaacacc agcgtgatca 2340
cccaqqcctq ccccaaggtg agcttcgagc ccatccccaa ccactactgc gcccccgccg 2400
gcttcgccat cctgaagtgc aaggacaaga agttcaacgg caagggcccc tgcaccaacg 2460
tgagcaccgt gcagtgcacc cacggcatcc gccccgtggt gagcacccag ctgctgctga 2520
acqqtaqcct qqccqagqag gagqtggtga tccgcagcgc taacttcgcc gacaacgcca 2580
aggtgateat egtgeagetg aacgagageg tggagateaa etgeaceege getagegeee 2640
actgcaacct gagccgcgc aagtggaacg acacctgaa caagatcgtg atcaagctgc 2700
gcgagcagtt cggcaacaag accatcgtgt tcaagcacag cagcggcggc gaccccgaga 2760
togtgaccca cagettcaac tgcggcggcg agttetteta etgcaacagc getageetgc 2820
cctgccgcat caagcagate atcaacatgt ggcagaaggt gggcaaggce atgtacgccc 2880
ccccatcag cggccagatc cgctgcagca gcaacatcac cggcctgctg ctgacccgcg 2940
acggcggcaa cagcaacaac gagagcgaga tcttccgtcc gggcggcggc gacatgcgcg 3000
acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc ctgggcgtgg 3060
ccccaccaa qqccaaqctt accgtccagg cccgccagct gctgaqcqqc atcgtgcagc 3120
```

| 200200000  |            | ~~~~       |            |            |            | 2200 |
|------------|------------|------------|------------|------------|------------|------|
|            |            | gccatcgagg |            |            |            |      |
|            |            | gcccgcaccc |            |            |            |      |
| agcagctcga | gcagatctgg | aaccacacca | cctggatgga | gtgggaccgc | gagatcaaca | 3300 |
| actacaccag | cctgatccac | agcctgatcg | aggagagcca | gaaccagcag | gagaagaacg | 3360 |
| agcaggagct | gctggagctg | gacaagtggg | ccagcctgtg | gaactggttc | aacatcacca | 3420 |
|            |            | ctgttcatca |            |            |            |      |
|            |            | atctgaggat |            |            |            |      |
|            |            | cccgtgcctt |            |            |            |      |
|            |            | gaaattgcat |            |            |            |      |
|            |            |            |            |            |            |      |
|            |            | gacagcaagg |            |            |            |      |
|            |            | atgggtaccc |            |            |            |      |
|            |            | catccccttc |            |            |            |      |
|            |            | ataggacact |            |            |            |      |
| cacccgctaa | agtacttgga | geggtetete | cctccctcat | cagcccacca | aaccaaacct | 3960 |
| agcctccaag | agtgggaaga | aattaaagca | agataggcta | ttaagtgcag | agggagagaa | 4020 |
| aatgcctcca | acatgtgagg | aagtaatgag | agaaatcata | gaattttaag | gccatgattt | 4080 |
| aaggccatca | tggccttaat | cttccgcttc | ctcgctcact | gactcgctgc | gctcggtcgt | 4140 |
|            |            | cagctcactc |            |            |            |      |
|            |            | acatgtgagc |            |            |            |      |
|            |            | ttttccatag |            |            |            |      |
|            |            | ggcgaaaccc |            |            |            |      |
|            |            | gctctcctgt |            |            |            |      |
|            |            |            |            |            |            |      |
|            |            | gcgtggcgct |            |            |            |      |
|            |            | ccaagctggg |            |            |            |      |
|            |            | actatcgtct |            |            |            |      |
|            |            | gtaacaggat |            |            |            |      |
|            |            | ctaactacgg |            |            |            |      |
|            |            | ccttcggaaa |            |            |            |      |
| aaccaccgct | ggtagcggtg | gtttttttgt | ttgcaagcag | cagattacgc | gcagaaaaaa | 4860 |
| aggatctcaa | gaagatcctt | tgatcttttc | tacggggtct | gacgctcagt | ggaacgaaaa | 4920 |
| ctcacgttaa | gggattttgg | tcatgagatt | atcaaaaagg | atcttcacct | agatcctttt | 4980 |
| aaattaaaaa | tgaagtttta | aatcaatcta | aagtatatat | gagtaaactt | ggtctgacag | 5040 |
| ttaccaatgc | ttaatcagtg | aggcacctat | ctcagcgatc | tgtctatttc | gttcatccat | 5100 |
|            |            | gggggcgctg |            |            |            |      |
|            |            | ccccatcatc |            |            |            |      |
|            |            | accagttggt |            |            |            |      |
|            |            | gcgtgatctg |            |            |            |      |
|            |            | tcaagtcagc |            |            |            |      |
|            |            | ctcatcgage |            |            |            |      |
|            |            | ttgaaaaagc |            |            |            |      |
|            |            |            |            |            |            |      |
|            |            | aagatcctgg |            |            |            |      |
|            |            | cccctcgtca |            |            |            |      |
|            |            | tgagaatggc |            | _          |            |      |
|            |            | ctcgtcatca |            |            |            |      |
|            |            | gagacgaaat |            |            |            |      |
|            |            | gcgcaggaac |            |            |            |      |
| gaatcaggat | attcttctaa | tacctggaat | gctgttttcc | cggggatcgc | agtggtgagt | 5940 |
| aaccatgcat | catcaggagt | acggataaaa | tgcttgatgg | tcggaagagg | cataaattcc | 6000 |
| gtcagccagt | ttagtctgac | catctcatct | gtaacatcat | tggcaacgct | acctttgcca | 6060 |
|            |            | cgcatcgggc |            |            |            |      |
|            |            | agcccattta |            |            |            |      |
|            |            | agacgtttcc |            |            |            |      |
|            |            | cagttttatt |            |            |            |      |
|            |            | agacacaacg |            |            |            |      |
|            |            | catgagcgga |            |            |            |      |
| 35         | J          |            |            |            | J          |      |

caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta agaaaccatt 6480 attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg tc 6532

```
<210> 70
<211> 7181
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 70
tegegegettt eggtgatgac ggtgaaaace tetgacacat geageteeeg gagacggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eccetttgge 1080
tettatgeat getatactgt tittggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagaetgae aeggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg ageeetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga acaccatgcg 1920
cgtgaaggag aagtaccagc acctgtggcg ctggggctgg cgctggggca ccatgctgct 1980
gggcatgctg atgatctgca gcgccaccga gaagctgtgg gtgaccgtgt actacggcgt 2040
gecegtgtgg aaggaggeea ceaceacet getetgegee agegaegeea aggeetaega 2100
caccgaggtg cacaacgtgt gggccaccca cgcctgcgtg cccaccgacc ccaacccca 2160
ggaggtggtg ctggtgaacg tgaccgagaa cttcgacatg tggaagaacg acatggtgga 2220
gcagatgcac gaggacatca tcagcctgtg ggaccagagc ctgaagccct gcgtgaagct 2280
gacccccctg tgcgtgagcc tgaagtgcac cgacctgaag aacgacacca acaccaacag 2340
cagcagcggc cgcatgatca tggagaaggg cgagatcaag aactgcagct tcaacatcag 2400
caccagcate egeggeaagg tgeagaagga gtaegeette ttetacaage tggacateat 2460
```

|            |            |            | hs *       |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | gcgtgatcac |      |
|            |            |            |            |            | cccccgccgg |      |
| cttcaccatc | ctgaagtgca | aggacaagaa | gttcaacggc | aagggcccct | gcaccaacgt | 2640 |
| gagcaccgtg | cagtgcaccc | acggcatccg | ccccgtggtg | agcacccagc | tgctgctgaa | 2700 |
| cgaccgcctg | gccgaggagg | aggtggtgat | ccgcagcgcc | aacttcgccg | acaacgccaa | 2760 |
|            |            |            |            |            | ccaacaacaa |      |
|            |            |            |            |            | gcgagatcat |      |
|            |            |            |            |            | acaccctgaa |      |
|            |            |            |            |            | tcaagcacag |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | agttcttcta |      |
|            |            |            |            |            | gcaccgaggg |      |
|            |            |            |            |            | agcagatcat |      |
|            |            |            |            |            | gccagatccg |      |
|            |            |            |            |            | gcaacaacga |      |
|            |            |            |            |            | gcgagctgta |      |
|            |            |            |            |            | ccaagcgccg |      |
| cgtggtgcag | cgcgagaagc | gcgccgtggg | catcggcgcc | ctgttcctgg | gcttcctggg | 3480 |
| cgccgccggc | agcaccatgg | gcgccgccag | catgaccctg | accgtgcagg | cccgccagct | 3540 |
| gctgagcggc | atcgtgcagc | agcagaacaa | cctgctgcgc | gccatcgagg | cccagcagca | 3600 |
|            |            |            |            |            | tggccgtgga |      |
|            |            |            |            |            | agcttatctg |      |
|            |            |            |            |            | agatctggaa |      |
|            |            |            |            |            | tgatccacag |      |
|            |            |            |            |            | tggagctgga |      |
|            |            |            |            |            | acatcaagct |      |
|            |            |            |            |            | tgctgagcat |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | tgcccatccc |      |
|            |            |            |            |            | accgcgaccg |      |
|            |            |            |            |            | ctagttgcca |      |
|            |            |            |            |            | ccactcccac |      |
|            |            |            |            |            | gtcattctat |      |
|            |            |            |            |            | atagcaggca |      |
|            |            |            |            |            | ccggttcctc |      |
| ctgggccaga | aagaagcagg | cacatcccct | tctctgtgac | acaccctgtc | cacgcccctg | 4500 |
| gttcttagtt | ccagccccac | tcataggaca | ctcatagctc | aggagggctc | cgccttcaat | 4560 |
| cccacccgct | aaagtacttg | gagcggtctc | tecetecete | atcagcccac | caaaccaaac | 4620 |
| ctageeteca | agagtgggaa | gaaattaaag | caagataggc | tattaagtgc | agagggagag | 4680 |
|            |            |            |            |            | aggccatcat |      |
|            |            |            |            |            | cggctgcggc |      |
|            |            |            |            |            | ggggataacg |      |
|            |            |            |            |            | aaggccgcgt |      |
|            |            |            |            |            | cgacgctcaa |      |
|            |            |            |            |            | cctggaagct |      |
|            |            |            |            |            | gcctttctcc |      |
|            |            |            |            |            | tcggtgtagg |      |
|            |            |            |            |            | cgctgcgcct |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | ccactggcag |      |
|            |            |            |            |            | gagttcttga |      |
|            |            |            |            |            | gctctgctga |      |
|            |            |            |            |            | accaccgctg |      |
|            |            |            |            | -          | ggatctcaag |      |
|            |            |            |            |            | tcacgttaag |      |
|            |            |            |            |            | aattaaaaat |      |
|            |            |            |            |            | taccaatgct |      |
| taatcagtga | ggcacctatc | tcagcgatct | gtctatttcg | ttcatccata | gttgcctgac | 5760 |
|            |            |            |            |            |            |      |

```
tegggggggg ggggegetga ggtetgeete gtgaagaagg tgttgetgae teataceagg 5820
cetgaatege eccateatee agecagaaag tgagggagee aeggttgatg agagetttgt 5880
cgggaagatg cgtgatctga tccttcaact cagcaaaagt tcgatttatt caacaaagcc 6000
gccgtcccgt caagtcagcg taatgctctg ccagtgttac aaccaattaa ccaattctga 6060
ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat 6120
accatatttt tgaaaaagcc, gtttctgtaa tgaaggagaa aactcaccga ggcagttcca 6180
taggatggca agatectggt ateggtetge gatteegaet egteeaacat caatacaace 6240
tattaatttc ccctcgtcaa aaataaggtt atcaagtgag aaatcaccat gagtgacgac 6300
tgaatccggt gagaatggca aaagcttatg catttctttc cagacttgtt caacaggcca 6360
gccattacgc tcgtcatcaa aatcactcgc atcaaccaaa ccgttattca ttcgtgattg 6420
cgcctgagcg agacgaaata cgcgatcgct gttaaaagga caattacaaa caggaatcga 6480
atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata 6540
ttcttctaat acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcatc 6600
atcaggagta cggataaaat gcttgatggt cggaagaggc ataaattccg tcagccagtt 6660
tagtetgace ateteatetg taacateatt ggcaacgeta cetttgecat gtttcagaaa 6720
caactetgge gcateggget teccatacaa tegatagatt gtegeacetg attgeecgae 6780
attategega geceatttat acceatataa ateageatee atgttggaat ttaategegq 6840
cctcqaqcaa qacgtttccc gttqaatatg gctcataaca ccccttgtat tactqtttat 6900
gtaagcagac agttttattg ttcatgatga tatattttta tcttgtgcaa tgtaacatca 6960
gagattttga gacacaacgt ggctttcccc cccccccat tattgaagca tttatcaggg 7020
ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt 7080
tecgegeaca ttteccegaa aagtgeeace tgaegtetaa gaaaccatta ttateatgae 7140
attaacctat aaaaataggo gtatcacgag gccctttcgt c
                                                               7181
```

```
<210> 71
<211> 7493
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
```

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcgggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cqtqqataqc qqtttgactc acgggqattt ccaaqtctcc accccattqa 780 cgtcaatggg agtttgttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 toccogtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca ccccbttggc 1080 tottabgcat gotatactgt ttttggcttg gggcctatac acccccgctt cottatgcta 1140

|            | tatagcttag               |            |            |            |            |      |
|------------|--------------------------|------------|------------|------------|------------|------|
|            | gatactttcc               |            |            |            |            |      |
|            | atgccaatac               |            |            |            |            |      |
|            | ccatttatta               |            |            |            |            |      |
|            | attaaacata               |            |            |            |            |      |
|            | tctccggtag               |            |            |            |            |      |
|            | ggtcgctcgg               |            |            |            |            |      |
|            | ccaccaccac               |            |            |            |            |      |
|            | gtggagattg               |            |            |            |            |      |
|            | atgcaggcag               |            |            |            |            |      |
|            | tgttaacggt               |            |            |            |            |      |
|            | gacataatag               |            |            |            |            |      |
|            | gtcgtcgaca               |            |            |            |            |      |
|            | aagtaccagc               |            |            |            |            |      |
|            | atgatctgca               |            |            |            |            |      |
|            | aaggaggcca               |            |            |            |            |      |
| caccgaggtg | cacaacgtgt               | gggccaccca | cgcctgcgtg | cccaccgacc | ccaaccccca | 2160 |
|            | ctggtgaacg               |            |            |            |            |      |
| gcagatgcac | gaggacatca               | tcagcctgtg | ggaccagagc | ctgaagccct | gcgtgaagct | 2280 |
| gacccccctg | tgcgtgagcc               | tgaagtgcac | cgacctgaag | aacgacacca | acaccaacag | 2340 |
|            | cgcatgatca               |            |            |            |            |      |
|            | cgcggcaagg               |            |            |            |            |      |
|            | aacgacacca               |            |            |            |            |      |
|            | cccaaggtga               |            |            |            |            |      |
|            | ctgaagtgca               |            |            |            |            |      |
|            | cagtgcaccc               |            |            |            |            |      |
|            | gccgaggagg               |            |            |            |            |      |
|            | gtgcagctga               |            |            |            |            |      |
|            | agcatccaca               |            |            |            |            |      |
|            | cgccaggccc               |            |            |            |            |      |
|            | atcaagctgc               |            |            |            |            |      |
|            | gaccccgaga               |            |            |            |            |      |
|            | acccagetgt               |            |            |            |            |      |
|            | accgagggca               |            |            |            |            |      |
|            | cagaaggtgg               |            |            |            |            |      |
|            | aacatcaccg               |            |            |            |            |      |
|            | ttccgcctgg               |            |            |            |            |      |
|            | gtggtgaaga               |            |            |            |            |      |
|            | cgcgagaagc               |            |            |            |            |      |
|            | agcaccatgg               |            |            |            |            |      |
|            | atcgtgcagc               |            |            |            |            |      |
|            | ctgaccgtgt               |            |            |            |            |      |
|            | aaggaccagc               |            |            |            |            |      |
|            | gtgccctgga               |            |            |            |            |      |
|            | tggatggagt               |            |            |            | -          |      |
|            | gagagccaga               |            |            |            |            |      |
|            | agcctgtgga               |            |            |            |            |      |
|            | atcgtgggcg               |            |            |            |            |      |
|            | gtgcgccagg               |            |            |            |            |      |
|            | gaccgccccg               |            |            |            |            |      |
|            | ctggtgaacg<br>taccaccgcc |            |            |            |            |      |
|            | cgcggctggg               |            |            |            |            |      |
|            | aagaacagcg               |            |            |            |            |      |
|            | cgcgtgatcg               |            |            |            |            |      |
| gggcaccgac | cycycyatty               | aggragraca | 22202cccdc | cycyccatoc | gocacacece | 3230 |

|            | cgccagggcc |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | cagccatctg |            |            |            |            |      |
|            | actgtccttt |            |            |            |            |      |
|            | attctggggg |            |            |            |            |      |
|            | catgctgggg |            |            |            |            |      |
|            | tcctgggcca |            |            |            |            |      |
|            | tggttcttag |            |            |            |            |      |
| tccgccttca | atcccacccg | ctaaagtact | tggagcggtc | tctccctccc | tcatcagccc | 4920 |
|            | acctagcctc |            |            |            |            |      |
| gcagaggag  | agaaaatgcc | tccaacatgt | gaggaagtaa | tgagagaaat | catagaattt | 5040 |
| taaggccatc | atggccttaa | tcttccgctt | cctcgctcac | tgactcgctg | cgctcggtcg | 5100 |
| ttcggctgcg | gcgagcggta | tcagctcact | caaaggcggt | aatacggtta | tccacagaat | 5160 |
| caggggataa | cgcaggaaag | aacatgtgag | caaaaggcca | gcaaaaggcc | aggaaccgta | 5220 |
| aaaaggccgc | gttgctggcg | tttttccata | ggctccgccc | ccctgacgag | catcacaaaa | 5280 |
|            | aagtcagagg |            |            |            |            |      |
|            | ctccctcgtg |            |            |            |            |      |
|            | cccttcggga |            |            |            |            |      |
|            | ggtcgttcgc |            |            |            |            |      |
|            | cttatccggt |            |            |            |            |      |
|            | agcagccact |            |            |            |            |      |
|            | gaagtggtgg |            |            |            |            |      |
|            | gaagccagtt |            |            |            |            |      |
|            | tggtagcggt |            |            |            |            |      |
|            | agaagatcct |            |            |            |            |      |
|            | agggattttg |            |            |            |            |      |
|            | atgaagtttt |            |            |            |            |      |
|            | cttaatcagt |            |            |            |            |      |
|            | actcgggggg |            |            |            |            |      |
| actcatacca | ggcctgaatc | gccccatcat | ccaqccaqaa | agtgaggag  | ccacaattaa | 6180 |
|            | gttgtaggtg |            |            |            |            |      |
|            | gtcgggaaga |            |            |            |            |      |
|            | ccgccgtccc |            |            |            |            |      |
|            | gattagaaaa |            |            |            |            |      |
|            | ataccatatt |            |            |            |            |      |
|            | cataggatgg |            |            |            |            |      |
|            | cctattaatt |            |            |            |            |      |
|            | actgaatccg |            |            |            |            |      |
|            | cagccattac |            |            |            |            |      |
|            | tgcgcctgag |            |            |            |            |      |
|            | gaatgcaacc |            |            |            |            |      |
|            | tattcttcta |            |            |            |            |      |
|            | tcatcaggag |            |            |            |            |      |
|            | tttagtctga |            |            |            |            |      |
|            | aacaactctg |            |            |            | _          |      |
|            | acattatcgc |            |            |            |            |      |
|            | ggcctcgagc |            |            |            |            |      |
|            | atgtaagcag |            |            |            |            |      |
|            | cagagatttt |            |            |            |            |      |
| -          | ggttattgtc |            |            |            |            |      |
|            | gttccgcgca |            |            |            |            |      |
|            | acattaacct |            |            |            |            | 7493 |
|            |            |            |            | - 22       | J          |      |

<210> 72

<211> 7186

<212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 72 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tqacqqtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acgggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccqcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 tccccqtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tottatgcat gotatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattqqtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattqqctat atqccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cqcaqttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 agggggteat ggtcgctcgg cagetecttg etectaacag tggaggccag acttaggcac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920 qtqaaqqaqa agtaccaqca cctqtqqcqc tgggqctggc gctqqqqcac catqctqctq 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040 cccgtgtgga aggaggccac caccacctg ctctgcgcca gcgacgccaa ggcctacgac 2100 accqagqtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280 accecetty gegtgageet gaagtgeace gaeetgaaga acgaeaceaa caccaacage 2340 agcageggee geatgateat ggagaaggge gagateaaga actgeagett caacateage 2400 accagcatec geggeaaggt geagaaggag taegeettet tetacaaget ggacateate 2460 cccatcgaca acgacaccac cagctacage ctgaccaget gcaacaccag cgtgatcacc 2520 caqqcctqcc ccaaqgtgag cttcgagccc atccccaacc actactgcgc ccccgccggc 2580 ttcaccatcc tgaaqtgcaa caacaagacc ttcaacggca ccggcccctg caccaacgtg 2640 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700 ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760

accatcatcq tqcaqctgaa caccagcgtg gagatcaact gcacccgccc caacaacaac 2820

|   | acccgcaagc | gaattcgcat | ccagcgcggc | cccggccgcg | ccttcgtgac | catcggcaag | 2880 |
|---|------------|------------|------------|------------|------------|------------|------|
|   |            |            |            | atcagccgcg |            |            |      |
|   |            |            |            | ttcggcaaca |            |            |      |
|   |            |            |            | acccacagct |            |            |      |
|   |            |            |            | agcacctggt |            |            |      |
|   |            |            |            | accatcaccc |            |            |      |
|   |            |            |            | gccatgtacg |            |            |      |
|   |            |            |            | ctgctgaccc |            |            |      |
|   |            |            |            | ggcgacatgc |            |            |      |
|   |            |            |            | ccctgggcg  |            |            |      |
|   |            |            |            | gtgggcatcg |            |            |      |
|   |            |            |            | gccagcatga |            |            |      |
|   |            |            |            | aacaacctgc |            |            |      |
|   |            |            |            | atcaagcagc |            |            |      |
| • |            |            |            |            |            |            |      |
|   |            |            |            | ctgggcatct |            |            |      |
|   |            |            |            | agetggagea |            |            |      |
|   |            |            |            | cgcgagatca |            |            |      |
|   |            |            |            | caggagaaga |            |            |      |
|   |            |            |            | ttcaacatca |            |            |      |
|   |            |            |            | gtgggcctgc |            |            |      |
|   |            |            |            | ageceetga  |            |            |      |
|   |            |            |            | atcgaggagg |            |            |      |
|   |            |            |            | tgaggatcca |            |            |      |
|   |            |            |            | gtgccttcct |            |            |      |
|   | cccactgtcc | tttcctaata | aaatgaggaa | attgcatcgc | attgtctgag | taggtgtcat | 4320 |
|   |            |            |            | agcaaggggg |            |            |      |
|   |            |            |            | ggtacccagg |            |            |      |
|   |            |            |            | ccccttctct |            |            |      |
|   | ccctggttct | tagttccagc | cccactcata | ggacactcat | agctcaggag | ggctccgcct | 4560 |
|   |            |            |            | gtctctccct |            |            |      |
|   |            |            |            | taaagcaaga |            |            |      |
|   | gagagaaaat | gcctccaaca | tgtgaggaag | taatgagaga | aatcatagaa | ttttaaggcc | 4740 |
|   |            |            |            | cactgactcg |            |            |      |
|   |            |            |            | ggtaatacgg |            |            |      |
|   |            |            |            | ccagcaaaag |            |            |      |
|   |            |            |            | ccccctgac  |            |            |      |
|   |            |            |            | actataaaga |            |            |      |
|   |            |            |            | cctgccgctt |            |            |      |
|   |            |            |            | tagctcacgc |            |            |      |
|   | gtaggtcgtt | cgctccaagc | tgggctgtgt | gcacgaaccc | cccgttcagc | ccgaccgctg | 5220 |
|   | cgccttatcc | ggtaactatc | gtcttgagtc | caacccggta | agacacgact | tatcgccact | 5280 |
|   | ggcagcagcc | actggtaaca | ggattagcag | agcgaggtat | gtaggcggtg | ctacagagtt | 5340 |
|   |            |            |            | tagaagaaca |            |            |      |
|   |            |            |            | tggtagctct |            |            |      |
|   |            |            |            | gcagcagatt |            |            |      |
|   | tcaagaagat | cctttgatct | tttctacggg | gtctgacgct | cagtggaacg | aaaactcacg | 5580 |
|   | ttaagggatt | ttggtcatga | gattatcaaa | aaggatcttc | acctagatcc | ttttaaatta | 5640 |
|   | aaaatgaagt | tttaaatcaa | tctaaagtat | atatgagtaa | acttggtctg | acagttacca | 5700 |
|   |            |            |            | gatctgtcta |            |            |      |
|   | ctgactcggg | ggggggggc  | gctgaggtct | gcctcgtgaa | gaaggtgttg | ctgactcata | 5820 |
|   |            |            |            | gaaagtgagg |            |            |      |
|   |            |            |            | gaacttttgc |            |            |      |
|   | gttgtcggga | agatgcgtga | tctgatcctt | caactcagca | aaagttcgat | ttattcaaca | 6000 |
|   |            |            |            | ctctgccagt |            |            |      |
|   |            |            |            | tgaaactgca |            |            |      |
|   |            |            |            |            |            |            |      |

```
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 6180
ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 6240
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 6300
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 6360
ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 6420
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 6480
atogaatgca acoggogcag gaacactgco agogoatcaa caatatttto acotgaatca 6540
qqatattctt ctaatacctg gaatgctgtt ttcccgggga tcgcagtggt gagtaaccat 6600
qcatcatcaq qaqtacqqat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6660
caqtttaqtc tqaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6720
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6780
ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt ggaatttaat 6840
egeggeeteg ageaagaegt tteeegttga atatggetea taacacceet tgtattactg 6900
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6960
catcagagat tttgagacac aacgtggctt tcccccccc cccattattg aagcatttat 7020
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 7080
qqqqttccqc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 7140
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc
```

<210> 73 <211> 7003 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 73 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagcttgtct gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttqqcqqqtq tcqqqqctqq cttaactatq cqqcatcaqa qcaqattqta ctqaqaqtqc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960

tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatct 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttaca 1320 ggatggggtc ccattatta tttacaaatt cacataaca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccqqtag cgqcqqagct tccacatccq agccctggtc ccatgctcc 1500

- 171 -

|            | ggtcgctcgg |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | ccaccaccac |            |            |            |            |      |
|            | gtggagattg |            |            |            |            |      |
|            | atgcaggcag |            |            |            |            |      |
|            | tgttaacggt |            |            |            |            |      |
|            | gacataatag |            |            |            | _          |      |
|            | gtcgtcgaca |            |            |            |            |      |
|            | agtaccagca |            |            |            |            |      |
|            | tgatctgcag |            |            |            |            |      |
|            | aggaggccac |            |            |            |            |      |
|            | acaacgtgtg |            |            |            |            |      |
| gaggtggtgc | tggtgaacgt | gaccgagaac | ttcgacatgt | ggaagaacga | catggtggag | 2220 |
| cagatgcacg | aggacatcat | cagcctgtgg | gaccagagcc | tgaagccctg | cgtgaagctg | 2280 |
| acccccctgt | gcgtgagcct | gaagtgcacc | gacctgaaga | acgacaccaa | caccaacagc | 2340 |
| agcageggcc | gcatgatcat | ggagaagggc | gagatcaaga | actgcagctt | caacatcagc | 2400 |
| accageatec | gcggcaaggt | gcagaaggag | tacgccttct | tctacaagct | ggacatcatc | 2460 |
| cccatcgaca | acgacaccac | cagctacagc | ctgaccagct | gcaacaccag | cgtgatcacc | 2520 |
| caggcctgcc | ccaaggtgag | cttcgagccc | atccccaacc | actactgcgc | ccccgccggc | 2580 |
|            | tgaagtgcaa |            |            |            |            |      |
|            | agtgcaccca |            |            |            |            |      |
|            | ccgaggagga |            |            |            |            |      |
|            | tgcagctgaa |            |            |            |            |      |
|            | gaattcgcat |            |            |            |            |      |
|            | tgcgccaggc |            |            |            |            |      |
|            | acagcaagct |            |            |            |            |      |
|            | geggegaeee |            |            |            |            |      |
|            | acagcaccca |            |            |            |            |      |
|            | acaacaccga |            |            |            |            |      |
|            | tgtggcagaa |            |            |            |            |      |
|            | gcagcaacat |            |            |            |            |      |
|            | agatetteeg |            |            |            |            |      |
|            | acaaggtggt |            |            |            |            |      |
|            | aggcccgcca |            |            |            |            |      |
|            | aggcccagca |            |            |            |            |      |
|            | tcctggccgt |            |            |            |            |      |
|            | cctggatgga |            |            |            |            |      |
|            | aggagagcca |            |            |            |            |      |
|            | ccagcctgtg |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | tgatcgtggg |            |            |            |            |      |
|            | gcgtgcgcca |            |            |            |            |      |
|            | ccgaccgccc |            |            |            |            |      |
|            | gcctggtgaa |            |            |            |            |      |
|            | ttgtttgccc |            |            |            |            |      |
|            | cctaataaaa |            |            |            |            |      |
|            | gtggggtggg |            |            |            |            |      |
|            | atgcggtggg |            |            |            |            |      |
|            | gaaagaagca |            |            |            |            |      |
|            | ttccagcccc |            |            |            |            |      |
|            | ctaaagtact |            |            |            |            |      |
|            | caagagtggg |            |            |            |            |      |
|            | tccaacatgt |            |            |            |            |      |
|            | tcttccgctt |            |            |            |            |      |
|            | tcagctcact |            |            |            |            |      |
|            | aacatgtgag |            |            |            |            |      |
| gttgctggcg | tttttccata | ggctccgccc | ccctgacgag | catcacaaaa | atcgacgctc | 4800 |

WO 02/32943 PCT/US01/25721 .

```
aagtcagagg tggcgaaacc cgacaggact ataaaqatac caggcqtttc cccctqqaaq 4860
ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct 4920
cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta 4980
ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc gttcagcccg accgctgcgc 5040
cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactqqc 5100
agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt 5160
gaagtggtgg cctaactacg gctacactag aagaacagta tttggtatct gcgctctgct 5220
gaagccagtt acctteggaa aaagagttgg tagetettga teeggeaaac aaaccaeege 5280
tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca 5340
aqaaqateet ttgatetttt etaeggggte tgacqeteag tggaacqaaa acteacqtta 5400
agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa 5460
atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg 5520
cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg 5580
actogggggg gggggcqct gaggtctqcc tcqtqaaqaa qqtqttqctq actcatacca 5640
qqcctqaatc qccccatcat ccaqccaqaa aqtqaqqqaq ccacqqttqa tqaqaqcttt 5700
gttgtaggtg gaccagttgg tgattttgaa cttttgcttt gccacggaac ggtctgcgtt 5760
gtcgggaaga tgcgtgatct gatccttcaa ctcagcaaaa gttcgattta ttcaacaaag 5820
ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt acaaccaatt aaccaattct 5880
gattagaaaa actcatcgag catcaaatga aactgcaatt tattcatatc aggattatca 5940
ataccatatt tttgaaaaag ccgtttctgt aatgaaggag aaaactcacc gaggcagttc 6000
cataggatgg caagatcctg gtatcggtct gcgattccga ctcgtccaac atcaatacaa 6060
cctattaatt tcccctcgtc aaaaataagg ttatcaagtg agaaatcacc atgagtgacg 6120
actgaatccg gtgagaatgg caaaagctta tgcatttctt tccagacttg ttcaacaggc 6180
cagccattac gctcgtcatc aaaatcactc gcatcaacca aaccgttatt cattcgtgat 6240
tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaag gacaattaca aacaggaatc 6300
gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa tattttcacc tgaatcagga 6360
tattetteta atacetggaa tgetgtttte ceggggateg cagtggtgag taaccatgca 6420
tcatcaggag tacggataaa atgcttgatg gtcggaagag gcataaattc cgtcagccag 6480
tttagtctga ccatctcatc tgtaacatca ttggcaacgc tacctttgcc atgtttcaga 6540
aacaactetg gegeateggg etteceatac aategataga ttgtegeace tgattgeeeg 6600
acattatege gageceattt atacecatat aaateageat eeatgttgga atttaatege 6660
ggcctcgagc aagacgtttc ccgttgaata tggctcataa caccccttgt attactgttt 6720
atgtaagcag acagttttat tgttcatgat gatatatttt tatcttgtqc aatqtaacat 6780
cagagatttt gagacacaac gtggctttcc cccccccc attattgaag catttatcag 6840.
ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg 6900
gttccgcgca catttccccg aaaagtgcca cctgacgtct aagaaaccat tattatcatg 6960
acattaacct ataaaaatag gcgtatcacg aggccctttc gtc
```

```
<210> 74
<211> 7186
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 74
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgg tcagcggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
```

|            |            | catatatgga |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            | acgacccccg |            |            |            |      |
|            |            | ctttccattg |            |            |            |      |
|            |            | aagtgtatca |            |            |            |      |
| tgacggtaaa | tggcccgcct | ggcattatgc | ccagtacatg | accttatggg | actttcctac | 660  |
| ttggcagtac | atctacgtat | tagtcatcgc | tattaccatg | gtgatgcggt | tttggcagta | 720  |
|            |            | ggtttgactc |            |            |            |      |
|            |            | ggcaccaaaá |            |            |            |      |
|            |            | tgggcggtag |            |            |            |      |
|            |            | agatcgcctg |            |            |            |      |
|            |            | ccagcctccg |            |            |            |      |
|            |            | taagtaccgc |            |            |            |      |
|            |            | ttttggcttg |            |            |            |      |
| taggtgatgg | tatagettag | cctataggtg | tagattatta | accattatto | accactcccc | 1200 |
| tattogtoac | gatactttcc | attactaatc | cataacataa | ctctttgcca | caactatctc | 1260 |
|            |            | tctgtccttc |            |            |            |      |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacaccat | ccccatacc  | 1380 |
|            |            | gcgtgggatc |            |            |            |      |
|            |            | cggcggagct |            |            |            |      |
|            |            | cagctccttg |            |            |            |      |
|            |            | cagtgtgccg |            |            |            |      |
| gaaaatgagc | atagagatta | ggctcgcacg | actalogecy | atggaggagg | tagggggg   | 1600 |
| gaaaacgage | atacagacag | ctgagttgtt | geegacgeag | acggaagacc | caaggcagcg | 1000 |
|            |            |            |            |            |            |      |
|            |            | ggagggcagt |            |            |            |      |
|            |            | ctgacagact |            |            |            |      |
|            |            | cgtgtgatca |            |            |            |      |
|            |            | cctgtggcgc |            |            |            |      |
|            |            | cgccaccgag |            |            |            |      |
| cccgcgcgga | aggaggeeae | caccaccctg | eccegegeca | gcgacgccaa | ggcctacgac | 2100 |
| accgaggtge | acaacgegeg | ggccacccac | geetgegtge | ccaccgaccc | caacccccag | 2160 |
| gaggtggtge | cggccgacgt | gaccgagaac | ttegacatgt | ggaagaacga | catggtggag | 2220 |
|            |            | cagcctgtgg |            |            |            |      |
|            |            | gaagtgcacc |            |            |            |      |
|            |            | ggagaagggc |            |            |            |      |
| accageacce | geggeaagge | gcagaaggag | tacgeettet | tctacaagct | ggacatcatc | 2460 |
|            |            | cagctacagc |            |            |            |      |
|            |            | cttcgagccc |            |            |            |      |
|            |            | cgacaagacc |            |            |            |      |
| agcaccgtgc | agtgcaccca | cggcatccgc | cccgtggtga | gcacccagct | gctgctgaac | 2700 |
| agcagcctaa | ccgaggagga | ggtggtgatc | cgcagcgtcg | acttcaccga | caacgccaag | 2760 |
|            |            | caccacggtg |            |            |            |      |
| accegeaage | gaattcgcat | ccagcgcggc | cccggccgcg | ccttcgtgac | catcggcaag | 2880 |
| atcggcaaca | tgcgccaggc | ccactgcgat | atcagccgcg | ccaagtggga | caacaccctg | 2940 |
| aagcagatcg | acagcaagct | gcgcgagcag | ttcggcaacg | acaagaccat | catcttcaag | 3000 |
| cagagcagcg | gcggcgaccc | cgagatcgtg | acccacaget | tcaactgcgg | cggcgagttc | 3060 |
|            |            | gctgttcaac |            |            |            |      |
| gagggcagca | acaacaccga | gggcagcgac | accatcaccc | tgccctgccg | catcaagcag | 3180 |
| atcatcaaca | tgtggcagaa | ggtgggcaag | gccatgtacg | cccccccat  | cagcggccag | 3240 |
| atccgctgca | gcagcgatat | caccggcctg | ctgctgaccc | gcgacggcgg | caacagcaac | 3300 |
| aacgagagcg | agatcttccg | cctgggcggc | ggcgacatgc | gcgacaactg | gcgcagcgag | 3360 |
| ctgtacaagt | acaaggtggt | gaagatcgag | cccctgggcg | tggcccccac | caaggccaag | 3420 |
| cgccgcgtgg | tgcagcgcga | gaagcgcgcc | gtgggcatcg | gcgccctgtt | cctgggcttc | 3480 |
| ctgggcgccg | ccggcagcac | catgggcgcc | gccagcatga | ccctgaccgt | gcaggcccgc | 3540 |
| cagctgctga | gcggcatcgt | gcagcagcag | aacaacctgc | tgcgcgccat | cgaggcccag | 3600 |
| cagcacctgc | tgcagctgac | cgtgtggggc | atcaagcagc | tgcaggcccg | catcctggcc | 3660 |
|            |            |            |            |            |            |      |

|            | acctgaagga |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | ccgccgtgcc |            |            |            |            |      |
|            | ccacctggat |            |            |            |            |      |
|            | tcgaggagag |            |            |            |            |      |
| ctggacaagt | gggccagcct | gtggaactgg | ttcaacatca | ccaactggct | gtggtacatc | 3960 |
| aagctgttca | tcatgatcgt | gggcggcctg | gtgggcctgc | gcatcgtgtt | cgccgtgctg | 4020 |
| agcatcgtga | accgcgtgcg | ccagggctac | agccccctga | gcttccagac | ccacctgccc | 4080 |
| atcccccgcg | gccccgaccg | ccccgagggc | atcgaggagg | agggcggcga | gcgcgaccgc | 4140 |
| gaccgcacga | tccgcctggt | gaacggcagc | tgaggatcca | gatctgctgt | gccttctagt | 4200 |
| tgccagccat | ctgttgtttg | cccctcccc  | gtgccttcct | tgaccctgga | aggtgccact | 4260 |
| cccactgtcc | tttcctaata | aaatgaggaa | attgcatcgc | attgtctgag | taggtgtcat | 4320 |
| tctattctgg | ggggtggggt | ggggcaggac | agcaaggggg | aggattggga | agacaatagc | 4380 |
| aggcatgctg | gggatgcggt | gggctctatg | ggtacccagg | tgctgaagaa | ttgacccggt | 4440 |
| tcctcctggg | ccagaaagaa | gcaggcacat | ccccttctct | gtgacacacc | ctgtccacgc | 4500 |
| ccctggttct | tagttccagc | cccactcata | ggacactcat | agctcaggag | ggctccgcct | 4560 |
|            | ccgctaaagt |            |            |            |            |      |
| caaacctagc | ctccaagagt | gggaagaaat | taaagcaaga | taggctatta | agtgcagagg | 4680 |
| gagagaaaat | gcctccaaca | tgtgaggaag | taatgagaga | aatcatagaa | ttttaaggcc | 4740 |
| atcatggcct | taatcttccg | cttcctcgct | cactgactcg | ctgcgctcgg | tcgttcggct | 4800 |
| gcggcgagcg | gtatcagctc | actcaaaggc | ggtaatacgg | ttatccacag | aatcagggga | 4860 |
| taacgcagga | aagaacatgt | gagcaaaagg | ccagcaaaag | gccaggaacc | gtaaaaaggc | 4920 |
|            | gcgtttttcc |            |            |            |            |      |
| ctcaagtcag | aggtggcgaa | acccgacagg | actataaaga | taccaggcgt | ttccccctgg | 5040 |
|            | gtgcgctctc |            |            |            |            |      |
|            | ggaagcgtgg |            |            |            |            |      |
|            | cgctccaagc |            |            |            |            |      |
|            | ggtaactatc |            |            |            |            |      |
|            | actggtaaca |            |            |            |            |      |
|            | tggcctaact |            |            |            |            |      |
|            | gttaccttcg |            |            |            |            |      |
|            | ggtggttttt |            |            |            |            |      |
|            | cctttgatct |            |            |            |            |      |
|            | ttggtcatga |            |            |            |            |      |
|            | tttaaatcaa |            |            |            |            |      |
|            | agtgaggcac |            |            |            |            |      |
|            | ggggggggc  |            |            |            |            |      |
|            | atcgccccat |            |            |            |            |      |
|            | gtggaccagt |            |            |            |            |      |
| gttgtcggga | agatgcgtga | tctgatcctt | caactcagca | aaagttcgat | ttattcaaca | 6000 |
| aagccgccgt | cccgtcaagt | cagcgtaatg | ctctgccagt | gttacaacca | attaaccaat | 6060 |
|            | aaaactcatc |            |            |            |            |      |
| tcaataccat | atttttgaaa | aagccgtttc | tgtaatgaag | gagaaaactc | accgaggcag | 6180 |
|            | tggcaagatc |            |            |            |            |      |
|            | atttcccctc |            |            |            |            |      |
|            | ccggtgagaa |            |            |            |            |      |
| ggccagccat | tacgctcgtc | atcaaaatca | ctcgcatcaa | ccaaaccgtt | attcattcgt | 6420 |
| gattgcgcct | gagcgagacg | aaatacgcga | tcgctgttaa | aaggacaatt | acaaacagga | 6480 |
| atcgaatgca | accggcgcag | gaacactgcc | agcgcatcaa | caatattttc | acctgaatca | 6540 |
| ggatattctt | ctaatacctg | gaatgctgtt | ttcccgggga | tcgcagtggt | gagtaaccat | 6600 |
|            | gagtacggat |            |            |            |            |      |
| cagtttagtc | tgaccatctc | atctgtaaca | tcattggcaa | cgctaccttt | gccatgtttc | 6720 |
| agaaacaact | ctggcgcatc | gggcttccca | tacaatcgat | agattgtcgc | acctgattgc | 6780 |
|            | cgcgagccca |            |            |            |            |      |
| cgcggcctcg | agcaagacgt | ttcccgttga | atatggctca | taacacccct | tgtattactg | 6900 |
| tttatgtaag | cagacagttt | tattgttcat | gatgatatat | ttttatcttg | tgcaatgtaa | 6960 |

```
catcagagat titgagacac aacgiggott toccoccocc cocattatig aagcattiat 7020 cagggitati gictcatgag oggatacata titgaatgia titagaaaaa taaacaaata 7080 ggggitoogo goacattico oogaaaagig ocaccigacg totaagaaac cattatiato 7140 atgacattaa cotataaaaa taggogtato aogaggooct titogio 7186
```

<210> 75 <211> 7186 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 75 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagacqqtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acqcccccta ttgacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtacege etatagaete tataggeaca eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatetc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcqt tgctqccqcq 1800 cgcgccacca gacataatag ctgacagact aacagactgt tectttccat gggtcttttc 1860 tgcaqtcacc qtcqtcqaca cgtgtgatca gatatcqcqq ccqctctaqa caccatqcqc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcqtq 2040 cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacqccaa ggcctacqac 2100 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160 gaggtggtgc tggtcgacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280

acceccetgt gegtgageet gaagtgeace gacetgaaga acgacaceaa caccaacage 2340

|            |            | ggagaagggc |             |            |            |      |
|------------|------------|------------|-------------|------------|------------|------|
|            |            | gcagaaggag |             |            |            |      |
|            |            | cagctacagc |             |            |            |      |
|            |            | cttcgagccc |             |            |            |      |
|            |            | cgacaagacc |             |            |            |      |
|            |            | cggcatccgc |             |            |            |      |
|            |            | ggtggtgatc |             |            |            |      |
| accatcatcg | tgcaggtgga | caccacggtg | gagatcgatt  | gcacccgccc | caacaacaac | 2820 |
| acccgcaagc | gaattcgcat | ccagcgcggc | cccggccgcg  | ccttcgtgac | catcggcaag | 2880 |
|            |            | ccactgcaac |             |            |            |      |
|            |            | gcgcgagcag |             |            |            |      |
|            |            | cgagatcgtg |             |            |            |      |
|            |            | gctgttcaac |             |            |            |      |
|            |            | gggcagcgac |             |            |            |      |
|            |            | ggtgggcaag |             |            |            |      |
|            |            | caccggcctg |             |            |            |      |
|            |            | cctgggcggc |             |            |            |      |
|            |            | gaagatcgag |             |            |            |      |
|            |            | gaagegegee |             |            |            |      |
|            |            |            |             |            |            |      |
|            |            | catgggcgcc |             |            |            |      |
|            |            | gcagcagcag |             |            |            |      |
|            |            | cgtgtggggc |             |            |            |      |
|            |            | ccagcagctg |             |            |            |      |
| atctgcacca | ccgccgtgcc | ctggaacgcc | agctggagca  | acaagagcct | ggagcagatc | 3780 |
|            |            | ggagtgggac |             |            |            |      |
|            |            | ccagaaccag |             |            |            |      |
|            |            | gtggaactgg |             |            |            |      |
|            |            | gggcggcctg |             |            |            |      |
|            |            | ccagggctac |             |            |            |      |
|            |            | ccccgagggc |             |            |            |      |
| gaccgcacga | tccgcctggt | gaacggcagc | tgaggatcca  | gatctgctgt | gccttctagt | 4200 |
| tgccagccat | ctgttgtttg | cccctcccc  | gtgccttcct  | tgaccctgga | aggtgccact | 4260 |
| cccactgtcc | tttcctaata | aaatgaggaa | attgcatcgc  | attgtctgag | taggtgtcat | 4320 |
| tctattctgg | ggggtggggt | ggggcaggac | agcaaggggg  | aggattggga | agacaatagc | 4380 |
|            |            | gggctctatg |             |            |            |      |
|            |            | gcaggcacat |             |            |            |      |
| ccctggttct | tagttccagc | cccactcata | ggacactcat  | agctcaggag | ggctccgcct | 4560 |
| tcaatcccac | ccgctaaagt | acttggagcg | gtctctccct  | ccctcatcag | cccaccaaac | 4620 |
|            |            | gggaagaaat |             |            |            |      |
| gagagaaaat | gcctccaaca | tgtgaggaag | taatgagaga  | aatcatagaa | ttttaaggcc | 4740 |
| atcatqqcct | taatcttccq | cttcctcgct | cactgactcg  | ctacactcaa | tcattcaact | 4800 |
| gcggcgagcg | qtatcaqctc | actcaaaggc | ggtaatacgg  | ttatccacag | aatcagggga | 4860 |
|            |            | gagcaaaagg |             |            |            |      |
|            |            | ataggeteeg |             |            |            |      |
| ctcaagtcag | aggtagcaaa | acccgacagg | actataaaga  | taccaggggt | ttcccctaa  | 5040 |
|            |            | ctgttccgac |             |            |            |      |
| tetecettea | graagrotee | cgctttctca | tageteacac  | tatagatata | tgagttggt  | 5100 |
| ataggtcatt | cactccaacc | tgggctgtgt | acacassacc  | cccattaaca | ccaaccact~ | 2700 |
| carcttatee | ggtaactate | gtcttgagtc | Seachagaece | agacacach  | tatogooga  | 5240 |
| adcaccacc  | actooteacc | grattagee  | acceacetat  | abadagaet  | chacacach  | J200 |
| chtcaactac | toggetaaca | ggattagcag | tagaaaaaaa  | graggeggrg | tatagaget  | 2340 |
|            |            | acggctacac |             |            |            |      |
| guryaageda | greatetteg | gaaaaagagt | cygragetet  | cgatccggca | aacaaaccac | 5460 |
| bassesse   | ggeggeete. | ttgtttgcaa | gcagcagatt  | acgcgcagaa | aaaaaggatc | 5520 |
| ccaagaagat | cccccgatcc | tttctacggg | gcccgacgct  | cagtggaacg | aaaactcacg | 5580 |
| ccaagggatt | ccggccatga | gattatcaaa | aaggatette  | acctagatcc | ccccaaatta | 5640 |

```
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttqqtctq acagttacca 5700
atgettaatc agtgaggeac ctateteage gatetgteta tttegtteat ccatagttgc 5760
ctgactcggg ggggggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata 5820
ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5880
tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc 5940
gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca 6000
aageegeegt ceegteaagt cagegtaatg ctetgecagt gttacaacea attaaceaat 6060
totgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 6120
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 6180
ttccatagga tggcaagatc ctggtatcgg tctqcqattc cqactcqtcc aacatcaata 6240
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 6300
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 6360
ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 6420
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacaqqa 6480°
ategaatgca aceggegeag gaacactgce agegeateaa caatatttte acetgaatea 6540
ggatattett etaatacetg gaatgetgtt tteeegggga tegeagtggt gagtaaceat 6600
gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6660
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6720
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6780
ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt ggaatttaat 6840
cgcggcctcg agcaagacgt ttcccgttga atatggctca taacacccct tgtattactg 6900
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttq tqcaatqtaa 6960
catcagagat titgagacac aacgiggett teceeecee eccattatig aageatitat 7020
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 7080
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 7140
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc
```

```
<210> 76
<211> 7498
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta cogccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccqcqtt acataactta cqqtaaatqq 420
cocqcctqqc tqaccqccca acqacccccq cccattqacq tcaataatqa cqtatqttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaaqtctcc accccattqa 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac egggacegat ceagecteeg eggeegggaa eggtgeattg gaaegeggat 1020
```

| teccegtgee  | aagagtgacg                                 | taagtaccgc  | ctatagactc  | tataggcaca | cccctttggc | 1080 |
|-------------|--|-------------|-------------|------------|------------|------|
| tettatgeat  | gctatactgt                                 | ttttggcttg  | gggcctatac  | acccccgctt | ccttatgcta | 1140 |
| taggtgatgg  | tatagettag                                 | cctataggtg  | tgggttattg  | accattattg | accactcccc | 1200 |
| tattggtgac  | gatactttcc                                 | attactaatc  | cataacatgg  | ctctttgcca | caactatctc | 1260 |
| tattggctat  | atgccaatac                                 | tctgtccttc  | agagactgac  | acggactctg | tatttttaca | 1320 |
| ggatggggtc  | ccatttatta                                 | tttacaaatt  | cacatataca  | acaacgccgt | cccccgtgcc | 1380 |
| cgcagttttt  | attaaacata                                 | gcgtgggatc  | tccacgcgaa  | tctcgggtac | gtgttccgga | 1440 |
| catgggctct  | tctccggtag                                 | cggcggagct  | tccacatccg  | agccctggtc | ccatgcctcc | 1500 |
| agcggctcat  | ggtcgctcgg                                 | cagctccttg  | ctcctaacag  | tggaggccag | acttaggcac | 1560 |
| agcacaatgc  | ccaccaccac                                 | cagtgtgccg  | cacaaggccg  | tggcggtagg | gtatgtgtct | 1620 |
|             |  |             |             |            | taaggcagcg |      |
|             |  |             |             |            | ggtaactccc |      |
| gttgcggtgc  | tgttaacggt                                 | ggagggcagt  | gtagtctgag  | cagtactcgt | tgctgccgcg | 1800 |
| cgcgccacca  | gacataatag                                 | ctgacagact  | aacagactgt  | tcctttccat | gggtcttttc | 1860 |
| tgcagtcacc  | gtcgtcgaca                                 | cgtgtgatca  | gatatcgcgg  | ccgctctaga | caccatgcgc | 1920 |
|             |  |             |             |            | catgctgctg |      |
|             |  |             |             |            | ctacggcgtg |      |
| cccgtgtgga  | aggaggccac                                 | caccaccctg  | ctctgcgcca  | gcgacgccaa | ggcctacgac | 2100 |
| accgaggtgc  | acaacgtgtg                                 | ggccacccac  | gcctgcgtgc  | ccaccgaccc | caacccccag | 2160 |
| gaggtggtgc  | tggtgaacgt                                 | gaccgagaac  | ttcgacatgt  | ggaagaacga | catggtggag | 2220 |
| cagatgcacg  | aggacatcat                                 | cagcctgtgg  | gaccagagcc  | tgaagccctg | cgtgaagctg | 2280 |
| accccctgt   | gcgtgagcct                                 | gaagtgcacc  | gacctgaaga  | acqacaccaa | caccaacage | 2340 |
| agcagcggcc  | gcatgatcat                                 | ggagaagggc  | gagatcaaga  | actocaoctt | caacatcagc | 2400 |
| accagcatcc  | gcggcaaggt                                 | gcagaaqqaq  | tacqccttct  | tctacaagct | ggacatcatc | 2460 |
| cccatcgaca  | acqacaccac                                 | caqctacaqc  | ctgaccaget  | gcaacaccag | cgtgatcacc | 2520 |
| caggectgee  | ccaaqqtqaq                                 | cttcgagccc  | atccccaacc  | actactococ | eccegeegge | 2580 |
| ttcaccatcc  | tgaagtgcaa                                 | caacaagacc  | ttcaacggca  | ccaaccccta | caccaacgtg | 2640 |
| agcaccgtgc  | agtgcaccca                                 | caacatccac  | cccataataa  | gcacccagct | gctgctgaac | 2700 |
| ggcagcctgg  | ccgaggagga                                 | ggtggtgatc  | cacaacataa  | acttcaccga | caacgccaag | 2760 |
| accatcatcq  | tacaactaa                                  | caccagcata  | gagatcaact  | aceccacca  | caacaacaac | 2820 |
| acccgcaage  | gaattcgcat                                 | ccagcgcggc  | cccaaccaca  | ccttcatasc | catcggcaag | 2880 |
| atcggcaaca  | tacaccagac                                 | ccactgcaac  | atcagccgcg  | ccaagtggaa | caacaccctg | 2940 |
| aagcagatcg  | acagcaagct                                 | acacasacsa  | ttcggcaaca  | acaacaccat | catcttcaag | 2000 |
| cagagcagcg  | acaacaaccc                                 | caagatcata  | acceacacct  | traactoro  | cggcgagttc | 3060 |
| ttctactcca  | acagcaccca                                 | gctgttcaac  | agcacctggt  | tcaacagcac | ctggagcacc | 3130 |
| gagggcagca  | acaacaccga                                 | addcadcdac  | accatcaccc  | taccetacca | catcaagcag | 3180 |
| atcatcaaca  | tataacaaaa                                 | agtagasaa   | accatataca  | cccccccat  | caccaagcag | 3740 |
| atccactaca  | gcagcaacat                                 | caccaaccta  | ctactasacc  | acascacac  | caacagcaac | 3240 |
| aacgagagcg  | agatetteeg                                 | cctagacaac  | accascatac  | gegacggegg | gcgcagcgag | 3300 |
| ctgtacaagt  | acaaggtggt                                 | gaagatcgag  | ccctagaca   | tagecece   | caaggccaag | 3430 |
| caccacataa  | tacaacacaa                                 | gaageegeg   | atagggata   | acaccctatt | cctgggcttc | 3420 |
| ctagacacca  | ccaacaacac                                 | catagagagac | accadeatea  | ccctaaccat | gcaggcccgc | 3540 |
| cagctgctga  | acaacatcat                                 | acsacsacsa  | aacaacctgc  | tacacaccat | cgaggcccag | 3500 |
| cagcacctgc  | tacaactaac                                 | catatagaac  | atcaaccege  | tacaaaccaa | catcctggcc | 3660 |
| atagaacact  | acchasagas                                 | ccaccacta   | changeage   | cacagacaca | cggcaagctt | 3000 |
| atctocacca  | ccccataca                                  | ctageageeg  | aggiggaccc  | ggggctgcag | ggagcagatc | 3720 |
| ragaaccaca  | ccacctogat                                 | ggagtggg    | cacagagea   | acaagageet | ggagcagatc | 3/80 |
| cacagootoa  | transassassassassassassassassassassassassa | 22a2c22gac  | Caccacacaca | acaactacac | cagcctgatc | 3840 |
| ctagacaach  | addccadec+                                 | atagaactag  | ttcaacataa  | acyaycagga | gctgctggag | 3900 |
| aagctottos  | tratratect                                 | grayaactyg  | atacacatca  | ccaactggct | gtggtacatc | 3960 |
| ageateatea  | accochace                                  | gggcggcccg  | argageetge  | geategeget | cgccgtgctg | 4020 |
| atcccccc    | accycytycy                                 | ccagggccac  | ageeeeetga  | gettecagae | ccacctgccc | 4080 |
| daccoccided | tagastest                                  | cecegaggge  | accgaggagg  | agggcggcga | gcgcgaccgc | 4140 |
| gaccycacga  | teegeetggt                                 | gaacggcagc  | coggeeetga  | tctgggacga | cctgcgcagc | 4200 |
| agactactac  | ccayccacca                                 | cegeeegege  | gacctgctgc  | rgarcgrgac | ccgcatcgtg | 4260 |
| gagergergg  | accacacaa                                  | ccgggaggcc  | ocgaagtact  | ggtggaacct | gctgcagtac | 4320 |

| tggagccagg | agctgaagaa | cagcgccgtg | agcctgctga | acgccaccgc   | catcgccgtg | 4380   |
|------------|------------|------------|------------|--------------|------------|--------|
| gccgagggca | ccgaccgcgt | gatcgaggtg | gtgcagggcg | cctgccgcgc   | cateegeeac | 4440   |
| atcccccgcc | gcatccgcca | gggcctggag | cgcatcctgc | tgtgaggatc   | cagatctgct | 4500   |
| gtgccttcta | gttgccagcc | atctgttgtt | tgcccctccc | ccgtgccttc   | cttgaccctg | 4560   |
| gaaggtgcca | ctcccactgt | cctttcctaa | taaaatgagg | aaattgcatc   | gcattgtctg | 4620   |
| agtaggtgtc | attctattct | ggggggtggg | gtggggcagg | acagcaaggg   | ggaggattgg | 4680   |
| gaagacaata | gcaggcatgc | tggggatgcg | gtgggctcta | tgggtaccca   | ggtgctgaag | 4740   |
| aattgacccg | gttcctcctg | ggccagaaag | aagcaggcac | atccccttct   | ctgtgacaca | 4800   |
| ccctgtccac | gcccctggtt | cttagttcca | gccccactca | taggacactc   | atageteagg | 4860   |
| agggctccgc | cttcaatccc | acccgctaaa | gtacttggag | cggtctctcc   | ctccctcatc | 4920   |
| agcccaccaa | accaaaccta | gcctccaaga | gtgggaagaa | attaaagcaa   | gataggctat | 4980   |
| taagtgcaga | gggagagaaa | atgcctccaa | catgtgagga | agtaatgaga   | gaaatcatag | 5040   |
| aattttaagg | ccatcatggc | cttaatcttc | cgcttcctcg | ctcactgact   | cgctgcgctc | 5100   |
| ggtcgttcgg | ctgcggcgag | cggtatcagc | tcactcaaag | gcggtaatac   | ggttatccac | 5160   |
| agaatcaggg | gataacgcag | gaaagaacat | gtgagcaaaa | ggccagcaaa   | aggccaggaa | 5220   |
| ccgtaaaaag | gccgcgttgc | tggcgttttt | ccataggctc | cgcccccctg   | acgagcatca | 5280   |
| caaaaatcga | cgctcaagtc | agaggtggcg | aaacccgaca | ggactataaa   | gataccagge | 5340   |
| gtttccccct | ggaagctccc | tegtgegete | tectgttecg | accetgeege   | ttaccggata | 5400   |
| cctgtccgcc | tttctccctt | cgggaagcgt | ggcgctttct | catageteae   | qctqtaqqta | 5460   |
| tctcagttcg | gtgtaggtcg | ttcgctccaa | gctgggctgt | qtqcacqaac   | cccccattca | 5520   |
| gcccgaccgc | tgcgccttat | ccggtaacta | tcqtcttqaq | tccaacccgg   | taagacacga | 5580   |
| cttatcgcca | ctggcagcag | ccactggtaa | caggattage | agagggaggt   | atgtaggcgg | 5640   |
| tgctacagag | ttcttgaagt | ggtggcctaa | ctacqqctac | actagaagaa   | cagtatttag | 5700   |
| tatctqcqct | ctqctqaaqc | cagttacctt | cqqaaaaaqa | gttggtagct   | cttgatccgg | 5760   |
| caaacaaacc | accqctqqta | gcggtggttt | ttttatttac | aagcagcaga   | ttacqcqcaq | 5820   |
| aaaaaaaqqa | tctcaagaag | atcctttgat | cttttctacq | gggtctgacg   | ctcagtggaa | 5880   |
| cqaaaactca | cqttaaqqqa | ttttggtcat | gagattatca | aaaaggatct   | tcacctagat | 5940   |
| ccttttaaat | taaaaatgaa | gttttaaatc | aatctaaagt | atatatgagt   | aaacttagtc | 6000   |
| tgacagttac | caatgcttaa | tcagtgaggc | acctatctca | gcgatctgtc   | tatttcgttc | 6060   |
| atccatagtt | gcctgactcg | aaaaaaaaaa | gcgctgaggt | ctacctcata   | aagaaggtgt | 6120   |
| tgctgactca | taccaggeet | gaatcgcccc | atcatccage | cagaaagtga   | addaaccaca | 6180   |
| gttgatgaga | gctttgttgt | aggtggacca | attagtaatt | ttgaactttt   | getttgecae | 6240   |
| ggaacggtct | gcgttgtcgg | gaagatgcgt | gatctgatcc | ttcaactcag   | caaaagttcg | 6300   |
| atttattcaa | caaaqccqcc | gtcccgtcaa | gtcagcgtaa | tactetacea   | gtgttacaac | 6360   |
| caattaacca | attctgatta | gaaaaactca | tcgagcatca | aatgaaactg   | caatttattc | 6420   |
| atatcaggat | tatcaatacc | atatttttga | aaaagccgtt | tctgtaatga   | aggagaaaac | 6480   |
| tcaccgagge | agttccatag | gatggcaaga | tcctggtatc | gatctacaat   | tecgaetect | 6540   |
| ccaacatcaa | tacaacctat | taatttcccc | tcqtcaaaaa | taaggttatc   | aagtgagaaa | 6600   |
| tcaccatgag | tgacgactga | atccggtgag | aatqqcaaaa | acttatacat   | ttctttccag | 6660   |
| acttgttcaa | caggccagcc | attacgctcg | tcatcaaaat | cactcgcatc   | aaccaaaccg | 6720   |
| ttattcattc | qtqattqcqc | ctgagcgaga | cgaaatacgc | gatcgctgtt   | aaaaggacaa | 6780   |
| ttacaaacaq | gaatcgaatg | caaccggcgc | aggaacactg | ccagcocate   | aacaatattt | 6840   |
| tcacctgaat | caggatattc | ttctaatacc | tggaatgctg | ttttcccaaa   | gatcgcagtg | 6900   |
| gtgagtaacc | atgcatcatc | aggagtacgg | ataaaatgct | taataatcaa   | aagagggata | 6960   |
| aattccqtca | gccagtttag | tctgaccatc | tcatctotaa | catcattage   | aacactacct | 7020   |
| ttqccatqtt | tcagaaacaa | ctctggcgca | tegagettee | catacaatco   | atagattoto | 7020   |
| gcacctgatt | gcccgacatt | ategegagee | catttatacc | catataaatc   | accatccata | 7140   |
| ttggaattta | atcgcggcct | cgagcaagac | atttcccatt | gaatatggct   | cataacaccc | 7200   |
| cttgtattac | tatttatata | agcagacagt | tttattotto | atgatgatat   | attttatct  | 7260   |
| tgtgcaatot | aacatcagag | attttgagac | acaacataac | tttcccccc    | ccccattat  | 7230   |
| tgaagcattt | atcagggtta | ttgtctcatg | agcggataca | tatttgaatg   | tatttages  | 7320   |
| aataaacaaa | taggggttee | gcgcacattt | CCCCGaaaaa | taccacctas   | catctaadaa | 7440   |
| accattatta | tcatgacatt | aacctataaa | aataggcgta | traccaecorga | ctttcctc   | 7498   |
|            |            |            | 350500     | Jungayyee    | coccycc    | / 7 70 |

<210> 77 <211> 7315

```
<212> DNA
 <213> Artificial Sequence
 <223> plasmid pVR1012x/s containing HIV genes
 <400> 77
 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
 ctccgccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
 agetegttta qtqaaccgtc\agategectg gagacgccat ccacgetgtt ttgacctcca 960
 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080
. tottatgcat gotatactgt ttttggottg gggootatac accoccgott cottatgcta 1140
 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200
 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
 qqatqqqqtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560
 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
 qttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800.
 cqcqccacca gacataatag ctgacagact aacagactgt teetttecat gggtetttte 1860
 tgcaqtcacc qtcqtcqaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
 gtgaaggaga agtaccagca cetgtggege tggggetgge getggggeac catgetgetg 1980
 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
 cccqtqtqqa aqqaqqccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100
 accqaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
 gaggtggtgc tggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
 cagatgcacg aggacatcat cagcetgtgg gaccagagee tgaageeetg egtgaagetg 2280
 accccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
 agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
 accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
 cccatcgaca acgacaccac cagctacage ctgaccaget gcaacaccag cgtgatcacc 2520
 caggeetgee ccaaggtgag ettegageee atecceaace actactgege eccegeegge 2580
 ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggcccctg caccaacgtg 2640
 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700
```

|            |            |            | cgcagcgtga |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            | gagatcaact |            |            |      |
| acccgcaagc | gaattcgcat | ccagcgcggc | cccggccgcg | ccttcgtgac | catcggcaag | 2880 |
| atcggcaaca | tgcgccaggc | ccactgcaac | atcagccgcg | ccaagtggaa | caacaccctg | 2940 |
| aagcagatcg | acagcaagct | gcgcgagcag | ttcggcaaca | acaagaccat | catcttcaag | 3000 |
|            |            |            | acccacaget |            |            |      |
|            |            |            | agcacctggt |            |            |      |
|            |            |            | accatcaccc |            |            |      |
|            |            |            | gccatgtacg |            |            |      |
|            |            |            | ctgctgaccc |            |            |      |
|            |            |            | ggcgacatgc |            |            |      |
|            |            |            |            |            |            |      |
|            |            |            | cccctgggcg |            |            |      |
|            |            |            | ggcatcgtgc |            |            |      |
|            |            |            | cagctgaccg |            |            |      |
|            |            |            | ctgaaggacc |            |            |      |
|            |            |            | gagatcaaca |            |            |      |
|            |            |            | gagaagaacg |            |            |      |
|            |            |            | aacatcacca |            |            |      |
| ctgttcatca | tgatcgtggg | cggcctggtg | ggcctgcgca | tcgtgttcgc | cgtgctgagc | 3840 |
| atcgtgaacc | gcgtgcgcca | gggctacagc | cccctgagct | tccagaccca | cctgcccatc | 3900 |
| ccccgcggcc | ccgaccgccc | cgagggcatc | gaggaggagg | gcggcgagcg | cgaccgcgac | 3960 |
|            |            |            | gccctgatct |            |            |      |
|            |            |            | ctgctgctga |            |            |      |
|            |            |            | aagtactggt |            |            |      |
|            |            |            | ctgctgaacg |            |            |      |
|            |            |            | cagggcgcct |            |            |      |
|            |            |            | atcctgctgt |            |            |      |
|            |            |            | ccctcccccg |            |            |      |
|            |            |            |            |            |            |      |
|            |            |            | aatgaggaaa |            |            |      |
|            |            |            | gggcaggaca |            |            |      |
|            |            |            | ggctctatgg |            |            |      |
|            |            |            | caggcacatc |            |            |      |
|            |            |            | ccactcatag |            |            |      |
| geteegeett | caaccccacc | cgctaaagta | cttggagcgg | teteteete  | cctcatcagc | 4740 |
|            |            |            | ggaagaaatt |            |            |      |
|            |            |            | gtgaggaagt |            |            |      |
|            |            |            | ttcctcgctc |            |            |      |
|            |            |            | ctcaaaggcg |            |            |      |
| atcaggggat | aacgcaggaa | agaacatgtg | agcaaaaggc | cagcaaaagg | ccaggaaccg | 5040 |
| taaaaaggcc | gcgttgctgg | cgtttttcca | taggeteege | cccctgacg  | agçatcacaa | 5100 |
| aaatcgacgc | tcaagtcaga | ggtggcgaaa | cccgacagga | ctataaagat | accaggcgtt | 5160 |
| tccccctgga | agctccctcg | tgcgctctcc | tgttccgacc | ctgccgctta | ccggatacct | 5220 |
|            |            |            | gctttctcat |            |            |      |
|            |            |            | gggctgtgtg |            |            |      |
|            |            |            | tcttgagtcc |            |            |      |
|            |            |            | gattagcaga |            |            |      |
|            |            |            | cggctacact |            |            |      |
|            |            |            | aaaaagagtt |            |            |      |
| acaaaccacc | actagrage  | ataattttt  | tgtttgcaag | carcagatta |            | 5640 |
|            |            |            | ttctacgggg |            |            |      |
|            |            |            |            |            |            |      |
|            |            |            | attatcaaaa |            |            |      |
|            |            |            | ctaaagtata |            |            |      |
|            |            |            | tatctcagcg |            | _          |      |
|            |            |            | ctgaggtctg |            |            |      |
| tgactcatac | caggcctgaa | tegeceeate | atccagccag | aaagtgaggg | agccacggtt | 6000 |

```
gatgagaget ttgttgtagg tggaccagtt ggtgattttg aacttttgct ttgccacgga 6060
acggtctgcg ttgtcgggaa gatgcgtgat ctgatccttc aactcagcaa aagttcgatt 6120
tattcaacaa agccgccgtc ccgtcaagtc agcgtaatgc tctgccagtg ttacaaccaa 6180
ttaaccaatt ctgattagaa aaactcatcg agcatcaaat qaaactqcaa tttattcata 6240
tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg agaaaactca 6300
ccgaggcagt tccataggat ggcaagatcc tggtatcggt ctgcgattcc gactcgtcca 6360
acatcaatac aacctattaa tttcccctcg tcaaaaataa ggttatcaag tgagaaatca 6420
ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc tttccagact 6480
tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac caaaccqtta 6540
ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa aggacaatta 6600
caaacaggaa togaatgcaa coggogoagg aacactgcca gogoatcaac aatattttca 6660
cctgaatcag gatattette taatacctgg aatgetqttt teccqqqqat cqcaqtqqtq 6720
agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaaq aggcataaat 6780
tecgteagee agtitagtet gaccatetea tetgtaacat cattggeaac getacetttg 6840
ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata gattgtcgca 6900
cetgattgcc egacattate gegageeeat ttataceeat ataaateage atceatgttg 6960
gaatttaatc geggeetega geaagaegtt teeegttgaa tatggeteat aacaeeeett 7020
gtaltactgt tratgtaagc agacagtitt attgttcatg atgatatatt titatcttgt 7080
gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc ccattattga 7140
agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 7200
aaacaaatag gggttccgcg cacatttccc cgaaaaqtgc cacctqacqt ctaaqaaacc 7260
attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtc
```

```
<210> 78
<211> 7498
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegettt eggegatgae ggegaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cqqtaaatqq 420
cocgcetgge tgaccgccca acgacccccg cccattgacq tcaataatga cqtatqttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960
tagaagacac cgggaccgat ccaqcctccg cggccqqqaa cqqtqcattq qaacqcqqat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eccetttgge 1080
tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
```

tattggetat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggecag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcaqaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 qttqcqqtqc tqttaacqqt qqaqqqcaqt qtaqtctqaq caqtactcqt tgctqccqcq 1800 cqcqccacca qacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040 cccgtgtgga aggaggccac caccacctg ctctgcgcca gcgacgccaa ggcctacgac 2100 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160 gaggtggtgc tggtcgacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280 acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340 agcaqcqqcc qcatgatcat ggagaagggc gagatcaagg actgtagctt cgacatcagc 2400 accagcatec geggeaaggt geagaaggag taegeettet tetacaaget ggacateate 2460 cccatcgaca acgacaccac cagctacagc ctgaccagct gcgaaactag tgtgatcacc 2520 caggeetgee ceaaggtgag ettegageee ateceeaace actactgege eccegeegge 2580 ttcaccatcc tgaagtgcaa cgacaagacc ttcgacggca ccgggccctg caccgacgtg 2640 agcacegtge agtgeaceca eggeateege eeegtggtga geacecaget getgetgaac 2700 ggcagcetgg cegaggagga ggtggtgate egcagegteg aetteacega caaegecaag 2760 accatcatcg tgcaggtgga caccacggtg gagatcgatt gcacccgccc caacaacaac 2820 accegeaage gaattegeat ecagegegge eceggeegeg cettegtgae categgeaag 2880 ateggeaaca tgegecagge ecactgegat ateageegeg ceaagtggga caacaceetg 2940 aagcagateg acagcaaget gegegageag tteggeaacg acaagaceat catetteaag 3000 cagagcagcg geggegaccc egagategtg acceacaget teaactgegg eggegagtte 3060 ttctactgcg acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120 gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180 atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240 atcogctgca gcagcgatat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300 aacgagagcg agatetteeg eetgggegge ggegacatge gegacaactg gegeagegag 3360 ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggcccccac caaggccaag 3420 egeogegtgg tgeagegega gaagegegee gtgggeateg gegeeetgtt cetgggette 3480 ctqqqcqccq ccqqcaqcac catqqqcqcc gccagcatga ccctgaccgt gcaggcccgc 3540 cagetgetga geggeategt geageageag aacaacetge tgegegeeat egaggeeeag 3600 cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660 gtggagcgct acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720 atotgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780 tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3840 cacageetga tegaggagag ceagaaceag caggagaaga acgageagga getgetggag 3900 ctggacaagt gggccagcct gtggaactgg ttcaacatca ccaactggct gtggtacatc 3960 aagetgttca teatgategt gggegeetg gtgggeetge geategtgtt egeegtgetg 4020 agcategtga acegegtgeg ecagggetac ageceeetga gettecagae ecacetgeee 4080 atcoccegeg geoccqaceg ceeegaggge ategaggagg agggeggega gegegacege 4140 gacegeacga teegeetggt gaaeggeage etggeeetga tetgggaega eetgegeage 4200 etgtgcctgt teagetacea eegeetgege gaeetgetge tgategtgae eegeategtg 4260 gagetgetgg geegeegegg etgggaggee etgaagtact ggtggaacet getgeagtac 4320 tggagccagg agetgaagaa cagegcegtg agectgetga aegecaeege categeegtg 4380 gccgagggca ccgaccgcgt gatcgaggtg gtgcagggcg cctgccgcgc catccgccac 4440 atcccccgcc gcatccgcca gggcctggag cgcatcctgc tgtgaggatc cagatctgct 4500 gtgccttcta gttgccagcc atctgttgtt tgcccctccc ccgtgccttc cttgaccctg 4560

| gaaggtgcca | ctcccactgt | cctttcctaa | taaaatgagg | aaattgcatc | gcattgtctg | 4620 |
|------------|------------|------------|------------|------------|------------|------|
|            | attctattct |            |            |            |            |      |
| gaagacaata | gcaggcatgc | tggggatgcg | gtgggctcta | tgggtaccca | ggtgctgaag | 4740 |
| aattgacccg | gttcctcctg | ggccagaaag | aagcaggcac | atccccttct | ctgtgacaca | 4800 |
| ccctgtccac | gcccctggtt | cttagttcca | gccccactca | taggacactc | atagctcagg | 4860 |
|            | cttcaatccc |            |            |            |            |      |
|            | accaaaccta |            |            |            |            |      |
|            | gggagagaaa |            |            |            |            |      |
|            | ccatcatggc |            |            |            |            |      |
| ggtcgttcgg | ctgcggcgag | cggtatcagc | tcactcaaag | gcggtaatac | ggttatccac | 5160 |
|            | gataacgcag |            |            |            |            |      |
| ccgtaaaaag | gccgcgttgc | tggcgttttt | ccataggctc | egeceeeetg | acgagcatca | 5280 |
|            | cgctcaagtc |            |            |            |            |      |
|            | ggaagctccc |            |            |            |            |      |
|            | tttctccctt |            |            |            |            |      |
|            | gtgtaggtcg |            |            |            |            |      |
|            | tgcgccttat |            |            |            |            |      |
|            | ctggcagcag |            |            |            |            |      |
|            | ttcttgaagt |            |            |            |            |      |
|            | ctgctgaagc |            |            |            |            |      |
|            | accgctggta |            |            |            |            |      |
|            | tctcaagaag |            |            |            |            |      |
|            | cgttaaggga |            |            |            |            |      |
|            | taaaaatgaa |            |            |            |            |      |
|            | caatgcttaa |            |            |            |            |      |
|            | gcctgactcg |            |            |            |            |      |
|            | taccaggcct |            |            |            |            |      |
|            | gctttgttgt |            |            |            |            |      |
|            | gcgttgtcgg |            |            |            |            |      |
|            | caaagccgcc |            |            |            |            |      |
|            | attctgatta |            |            |            |            |      |
|            | tatcaatacc |            |            |            |            |      |
|            | agttccatag |            |            |            |            |      |
|            | tacaacctat |            |            |            |            |      |
|            | tgacgactga |            |            |            |            |      |
|            | caggccagcc |            |            |            |            |      |
| ttattcattc | gtgattgcgc | ctgagcgaga | cgaaatacgc | gatcgctgtt | aaaaggacaa | 6780 |
| ttacaaacag | gaatcgaatg | caaccggcgc | aggaacactg | ccagcgcatc | aacaatattt | 6840 |
| tcacctgaat | caggatattc | ttctaatacc | tggaatgctg | ttttcccggg | gatcgcagtg | 6900 |
|            | atgcatcatc |            |            |            |            |      |
|            | gccagtttag |            |            |            |            |      |
| ttgccatgtt | tcagaaacaa | ctctggcgca | tcgggcttcc | catacaatcg | atagattgtc | 7080 |
|            | gcccgacatt |            |            |            |            |      |
|            | atcgcggcct |            |            |            |            |      |
|            | tgtttatgta |            |            |            |            |      |
|            | aacatcagag |            |            |            |            |      |
|            | atcagggtta |            |            |            |            |      |
| aataaacaaa | taggggttcc | gcgcacattt | ccccgaaaag | tgccacctga | cgtctaagaa | 7440 |
| accattatta | tcatgacatt | aacctataaa | aataggcgta | tcacgaggcc | ctttcgtc   | 7498 |
|            |            |            |            |            |            |      |

<210> 79

<211> 7498

<212> DNA

<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 79 tegegegttt eggtgatgae ggtgaaaacc tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cqtcaatqqq aqtttqtttt qqcaccaaaa tcaacqqqac tttccaaaat qtcqtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegitta gigaaccgic agategeetg gagacgeeat ceacgetqit tiqaccicea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgo ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980 ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040 cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160 gaggtggtgc tggtcgacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220 cagatgcacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaagetg 2280 accccctgt gcgtgagcct gaagtgcacc gacctgaaga acqacaccaa caccaacaqc 2340 agcagoggcc gcatgateat ggagaagggc gagateaagg actgtagett cgacateage 2400 accagcated geggeaaggt geagaaggag tacgeettet tetacaaget ggacateate 2460 cccategaca acgacaccae cagetacage etgaccaget gegaaactag tgtgatcace 2520 caggectgec ccaaggtgag ettegageec atececaace actaetgege eccegeegge 2580 ttcaccatcc tgaagtgcaa cgacaagacc ttcgacggca ccgggccctg caccgacgtg 2640 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700 ggcagcetgg ccgaggagga ggtggtgate egcagcgteg aetteacega caacgecaag 2760 accatcatcg tgcaggtgga caccacggtg gagatcgatt gcacccgccc caacaacaac 2820 accegeaage gaattegeat ceagegegge eeeggeegeg cettegtgae eateggeaag 2880 ateggeaaca tgegeeagge ceaetgeaac ateageegeg ceaagtggaa caacaceetg 2940

| aagcagatcg | acagcaagct | gcgcgagcag | ttcggcaaca | acaagaccat | catcttcaag | 3000 |
|------------|------------|------------|------------|------------|------------|------|
| cagagcagcg | gcggcgaccc | cgagatcgtg | acccacaget | tcaactgcgg | cggcgagttc | 3060 |
| ttctactgca | acagcaccca | gctgttcaac | agcacctggt | tcaacagcac | ctggagcacc | 3120 |
| gagggcagca | acaacaccga | gggcagcgac | accatcaccc | tgccctgccg | catcaagcag | 3180 |
| atcatcaaca | tgtggcagaa | ggtgggcaag | gccatgtacg | cccccccat  | cagcggccag | 3240 |
| atccgctgca | gcagcaacat | caccggcctg | ctgctgaccc | gcgacggcgg | caacagcaac | 3300 |
| aacgagagcg | agatcttccg | cctgggcggc | ggcgacatgc | gcgacaactg | gcgcagcgag | 3360 |
| ctgtacaagt | acaaggtggt | gaagatcgag | cccctgggcg | tggcccccac | caaggccaag | 3420 |
| cgccgcgtgg | tgcagcgcga | gaagcgcgcc | gtgggcatcg | gcgccctgtt | cctgggcttc | 3480 |
| ctgggcgccg | ccggcagcac | catgggcgcc | gccagcatga | ccctgaccgt | gcaggcccgc | 3540 |
| cagctgctga | gcggcatcgt | gcagcagcag | aacaacctgc | tgcgcgccat | cgaggcccag | 3600 |
| cagcacctgc | tgcagctgac | cgtgtggggc | atcaagcagc | tgcaggcccg | catcctggcc | 3660 |
| gtggagcgct | acctgaagga | ccagcagctg | ctgggcatct | ggggctgcag | cggcaagctt | 3720 |
| atctgcacca | ccgccgtgcc | ctggaacgcc | agctggagca | acaagagcct | ggagcagatc | 3780 |
| tggaaccaca | ccacctggat | ggagtgggac | cgcgagatca | acaactacac | cagcctgatc | 3840 |
| cacagcctga | tcgaggagag | ccagaaccag | caggagaaga | acgagcagga | gctgctggag | 3900 |
|            | gggccagcct |            |            |            |            |      |
| aagctgttca | tcatgatcgt | gggcggcctg | gtgggcctgc | gcatcgtgtt | cgccgtgctg | 4020 |
|            | accgcgtgcg |            |            |            |            |      |
|            | gccccgaccg |            |            |            |            |      |
| gaccgcacga | tccgcctggt | gaacggcagc | ctggccctga | tctgggacga | cctgcgcagc | 4200 |
| ctgtgcctgt | tcagctacca | ccgcctgcgc | gacctgctgc | tgatcgtgac | ccgcatcgtg | 4260 |
|            | gccgccgcgg |            |            |            |            |      |
|            | agctgaagaa |            |            |            |            |      |
|            | ccgaccgcgt |            |            |            |            |      |
|            | gcatccgcca |            |            |            |            |      |
|            | gttgccagcc |            |            |            |            |      |
|            | ctcccactgt |            |            |            |            |      |
|            | attctattct |            |            |            |            |      |
|            | gcaggcatgc |            |            |            |            |      |
|            | gttcctcctg |            |            |            |            |      |
|            | gcccctggtt |            |            |            |            |      |
| agggctccgc | cttcaatccc | acccgctaaa | gtacttggag | cggtctctcc | ctccctcatc | 4920 |
|            | accaaaccta |            |            |            |            |      |
|            | gggagagaaa |            |            |            |            |      |
| aattttaagg | ccatcatggc | cttaatcttc | cgcttcctcg | ctcactgact | cgctgcgctc | 5100 |
| ggtcgttcgg | ctgcggcgag | cggtatcagc | tcactcaaag | gcggtaatac | ggttatccac | 5160 |
| agaatcaggg | gataacgcag | gaaagaacat | gtgagcaaaa | ggccagcaaa | aggccaggaa | 5220 |
| ccgtaaaaag | gccgcgttgc | tggcgttttt | ccataggctc | cgccccctg  | acgagcatca | 5280 |
|            | cgctcaagtc |            |            |            |            |      |
|            | ggaagctccc |            |            |            |            |      |
|            | tttctccctt |            |            |            |            |      |
| tctcagttcg | gtgtaggtcg | ttcgctccaa | gctgggctgt | gtgcacgaac | cccccgttca | 5520 |
| gcccgaccgc | tgcgccttat | ccggtaacta | tcgtcttgag | tccaacccgg | taagacacga | 5580 |
| cttatcgcca | ctggcagcag | ccactggtaa | caggattagc | agagcgaggt | atgtaggcgg | 5640 |
| tgctacagag | ttcttgaagt | ggtggcctaa | ctacggctac | actagaagaa | cagtatttgg | 5700 |
|            | ctgctgaagc |            |            |            |            |      |
|            | accgctggta |            |            |            |            |      |
|            | tctcaagaag |            |            |            |            |      |
| cgaaaactca | cgttaaggga | ttttggtcat | gagattatca | aaaaggatct | tcacctagat | 5940 |
|            | taaaaatgaa |            |            |            |            |      |
|            | caatgcttaa |            |            |            |            |      |
| atccatagtt | gcctgactcg | 999999999  | gcgctgaggt | ctgcctcgtg | aagaaggtgt | 6120 |
| tgctgactca | taccaggcct | gaatcgcccc | atcatccagc | cagaaagtga | gggagccacg | 6180 |
| gttgatgaga | gctttgttgt | aggtggacca | gttggtgatt | ttgaactttt | gctttgccac | 6240 |

```
ggaacggtet gegttgtegg gaagatgegt gatetgatee tteaacteag caaaagtteg 6300
atttattcaa caaagccqcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 6360
caattaacca attotgatta gaaaaactca togagcatca aatgaaactg caatttatto 6420
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 6480
teacegagge agtteeatag gatggeaaga teetggtate ggtetgegat teegactegt 6540
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 6600
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 6660
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 6720
ttattcattc qtqattqcqc ctqaqcqaqa cqaaatacqc qatcqctqtt aaaaqqacaa 6780
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 6840
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg gatcgcaqtg 6900
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6960
aattoogtoa gocagtttag totgaccato toatotgtaa catcattggc aacgctacot 7020
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 7080
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 7140
ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 7200
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 7260
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc cccccattat 7320
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 7380
aataaacaaa taggggttee gegeacattt eecegaaaag tgecaeetga egtetaagaa 7440
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 7498
<210> 80
<211> 5549
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 80
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtct gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
```

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 eteegeeeca ttgacgeaaa tgggeggtag gegtgtaegg tgggaggtet atataageag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcetecg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgacg taagtacege ctatagacte tataggeaca eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cetataggtg tgggttattg accattattg accactecec 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320

|            | ccatttatta |             |            |            |            |      |
|------------|------------|-------------|------------|------------|------------|------|
| cgcagttttt | attaaacata | gcgtgggatc  | tccacgcgaa | tctcgggtac | gtgttccgga | 1440 |
|            | tctccggtag |             |            |            |            |      |
|            | ggtcgctcgg |             |            |            |            |      |
|            | ccaccaccac |             |            |            |            |      |
| gaaaatgagc | gtggagattg | ggctcgcacg  | gctgacgcag | atggaagact | taaggcagcg | 1680 |
|            | atgcaggcag |             |            |            |            |      |
| gttgcggtgc | tgttaacggt | ggagggcagt  | gtagtctgag | cagtactcgt | tgctgccgcg | 1800 |
| cgcgccacca | gacataatag | ctgacagact  | aacagactgt | tcctttccat | gggtcttttc | 1860 |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca  | gatatcgcgg | ccgctctaga | ccaccatggg | 1920 |
| cggcaagtgg | agcaagagca | gcgtgatcgg  | ctggcccgcc | gtgcgcgagc | gcatgcgccg | 1980 |
| cgccgagccc | gccgccgacg | gcgtgggcgc  | cgtgagccgc | gacctggaga | agcacggcgc | 2040 |
| catcaccagc | agcaacaccg | ccgccaacaa  | cgccgcctgc | gcctggctgg | aggcccagga | 2100 |
|            | gtgggcttcc |             |            |            |            |      |
|            | gacctgagcc |             |            |            |            |      |
| cagccagcgc | cgccaggaca | tcctggacct  | gtggatctac | cacacccagg | gctacttccc | 2280 |
|            | aactacaccc |             |            |            |            |      |
|            | gtgcccgtgg |             |            |            |            |      |
|            | caccccgtga |             |            |            |            |      |
|            | gacagccgcc |             |            |            |            |      |
|            | tgctgaggat |             |            |            |            |      |
|            | cccgtgcctt |             |            |            |            |      |
|            | gaaattgcat |             |            |            |            |      |
|            | gacagcaagg |             |            |            |            |      |
|            | atgggtaccc |             |            |            |            |      |
|            | cateceette |             |            |            |            |      |
|            | ataggacact |             |            |            |            |      |
|            | gcggtctctc |             |            |            |            |      |
|            | aattaaagca |             |            |            |            |      |
|            | aagtaatgag |             |            |            |            |      |
|            | gctcactgac |             |            |            |            |      |
| ctcactcaaa | ggcggtaata | cogttatcca  | cagaatcagg | goataacoca | acaacaaca  | 3240 |
| totoaccaaa | aggccagcaa | aaggccagga  | accotaaaaa | gaccacatta | ctacatttt  | 3300 |
|            | ccgccccct  |             |            |            |            |      |
|            | aggactataa |             |            |            |            |      |
|            | gaccctgccg |             |            |            |            |      |
|            | tcatagetca |             |            |            |            |      |
|            | tgtgcacgaa |             |            |            |            |      |
|            | gtccaacccg |             |            |            |            |      |
|            |            |             |            |            |            |      |
|            | cagagcgagg |             |            |            |            |      |
|            | cactagaaga |             |            |            |            |      |
|            | agttggtagc |             |            |            |            |      |
|            | caagcagcag |             |            |            |            |      |
|            | ggggtctgac |             |            |            |            |      |
| tgagattate | aaaaaggatc | tccacctaga  | tecttttaaa | ttaaaaatga | agttttaaat | 4020 |
| caacccaaag | tatatatgag | chalactegge | ctgacagtta | ccaatgetta | accagegagg | 4080 |
|            | agcgatctgt |             |            |            |            |      |
| ggegetgagg | tctgcctcgt | yaayaaggtg  | redecaree  | acaccaggcc | tgaatcgccc | 4200 |
| catcatccag | ccagaaagtg | agggagccac  | ggttgatgag | agctttgttg | taggtggacc | 4260 |
| agttggtgat | tttgaacttt | tgctttgcca  | cggaacggtc | tgcgttgtcg | ggaagatgcg | 4320 |
| rgatctgatc | cttcaactca | gcaaaagttc  | gatttattca | acaaagccgc | cgtcccgtca | 4380 |
| agtcagcgta | atgctctgcc | agtgttacaa  | ccaattaacc | aattctgatt | agaaaaactc | 4440 |
| accgagcatc | aaatgaaact | gcaatttatt  | catatcagga | ttatcaatac | catatttttg | 4500 |
| aaaaagccgt | ttctgtaatg | aaggagaaaa  | cccaccgagg | cagttccata | ggatggcaag | 4560 |
| accctggtat | cggtctgcga | ttccgactcg  | tccaacatca | atacaaccta | ttaatttccc | 4620 |

```
ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga 4680
gaatggcaaa agcttatgca tttctttcca gacttgttca acaggccagc cattacgctc 4740
gtcatcaaaa tcactcgcat caaccaaacc gttattcatt cgtgattgcg cctgagcgag 4800
acgaaatacg cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg 4860
caggaacact gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac 4920
ctggaatget gttttcccgg ggatcgcagt ggtgagtaac catgcatcat caggagtacg 4980
qataaaatqc ttqatqqtcq qaaqaqqcat aaattccqtc aqccaqttta gtctgaccat 5040
ctcatctqta acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc 5100
atogggotto coatacaato gatagattgt ogcacotgat tgccogacat tatogcgago 5160
ccatttatac ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga 5220
cqtttcccqt tqaatatqqc tcataacacc ccttgtatta ctqtttatqt aagcagacaq 5280
ttttattgtt catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga 5340
cacaacqtqq ctttcccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 5400
qaqqqatac atatttqaat qtatttaqaa aaataaacaa ataqqqqttc cqcqcacatt 5460
teccequaaa qtqceacetq acqtetaaqa aaccattatt atcatgacat taacctataa 5520
aaataggcgt atcacgaggc cctttcgtc
```

<210> 81 <211> 5549 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttqqcqgqtq tcgggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cqtcaatggq agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 ageteqttta qtqaaccqtc agategeetg gagaegeeat ccaegetgtt ttgaccteca 960 tagaagacac egggacegat ceageeteeg eggeegggaa eggtgeattg gaacgeggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140

taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactccc 1200 tattggtgac gatacttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacaccgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620

|            |            |            | gctgacgcag |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
|            |            |            | gtagtctgag |            |            |      |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | tcctttccat | gggtcttttc | 1860 |
|            |            |            | gatatcgcgg |            |            |      |
| cggcaagtgg | agcaagagca | gcgtgatcgg | ctggcccgcc | gtgcgcgagc | gcatgcgccg | 1980 |
| cgccgagccc | gccgccgacg | gcgtgggcgc | cgtgagccgc | gacctggaga | agcacggcgc | 2040 |
| catcaccagc | agcaacaccg | ccgccaacaa | cgccgcctgc | gcctggctgg | aggcccagga | 2100 |
|            |            |            | acaggtggca |            |            |      |
| ggccgccgtg | gacctgagcc | acttcctgaa | ggagaagggc | ggcctggagg | gcctgatcca | 2220 |
| cagccagcgc | cgccaggaca | tcctggacct | gtggatctac | cacacccagg | gctacttccc | 2280 |
| cgactggcag | aactacaccc | ccggccccgg | cgtgcgctac | cccctgacct | tcggctggtg | 2340 |
| ctacaagctg | gtgcccgtgg | agcccgacaa | ggtggaggag | gccaacaagg | gcgagaacac | 2400 |
| cagcctgctg | caccccgtga | gcctgcacgg | catggcagca | cccgagcgcg | aggtgctgga | 2460 |
| gtggcgcttc | gacagccgcc | tggccttcca | ccacgtggcc | cgcgagctgc | accccgagta | 2520 |
| cttcaagaac | tgctgaggat | ccagatctgc | tgtgccttct | agttgccagc | catctgttgt | 2580 |
|            |            |            | ggaaggtgcc |            |            |      |
|            |            |            | gagtaggtgt |            |            |      |
| ggtggggcag | gacagcaagg | gggaggattg | ggaagacaat | agcaggcatg | ctggggatgc | 2760 |
|            |            |            | gaattgaccc |            |            |      |
| gaagcaggca | catccccttc | tctgtgacac | accctgtcca | cgcccctggt | tcttagttcc | 2880 |
| agccccactc | ataggacact | catageteag | gagggctccg | ccttcaatcc | cacccgctaa | 2940 |
| agtacttgga | gcggtctctc | cctccctcat | cagcccacca | aaccaaacct | agcctccaag | 3000 |
| agtgggaaga | aattaaagca | agataggcta | ttaagtgcag | agggagagaa | aatgcctcca | 3060 |
|            |            |            | gaattttaag |            |            |      |
|            |            |            | cggtcgttcg |            |            |      |
| ctcactcaaa | ggcggtaata | cggttatcca | cagaatcagg | ggataacgca | ggaaagaaca | 3240 |
|            |            |            | accgtaaaaa |            |            |      |
|            |            |            | acaaaaatcg |            |            |      |
| gaaacccgac | aggactataa | agataccagg | cgtttccccc | tggaagctcc | ctcgtgcgct | 3420 |
| ctcctgttcc | gaccctgccg | cttaccggat | acctgtccgc | cttctccct  | tcgggaagcg | 3480 |
|            |            |            | atctcagttc |            |            |      |
| agctgggctg | tgtgcacgaa | cccccgttc  | agcccgaccg | ctgcgcctta | tccggtaact | 3600 |
| atcgtcttga | gtccaacccg | gtaagacacg | acttatcgcc | actggcagca | gccactggta | 3660 |
|            |            |            | gtgctacaga |            |            |      |
|            |            |            | gtatctgcgc |            |            |      |
|            |            |            | gcaaacaaac |            |            |      |
| tttttgtttg | caagcagcag | attacgcgca | gaaaaaaagg | atctcaagaa | gatcctttga | 3900 |
|            |            |            | acgaaaactc |            |            |      |
| tgagattatc | aaaaaggatc | ttcacctaga | tccttttaaa | ttaaaaatga | agttttaaat | 4020 |
| caatctaaag | tatatatgag | taaacttggt | ctgacagtta | ccaatgctta | atcagtgagg | 4080 |
| cacctatctc | agcgatctgt | ctatttcgtt | catccatagt | tgcctgactc | 999999999  | 4140 |
|            |            |            | ttgctgactc |            |            |      |
| catcatccag | ccagaaagtg | agggagccac | ggttgatgag | agctttgttg | taggtggacc | 4260 |
| agttggtgat | tttgaacttt | tgctttgcca | cggaacggtc | tgcgttgtcg | ggaagatgcg | 4320 |
| tgatctgatc | cttcaactca | gcaaaagttc | gatttattca | acaaagccgc | cgtcccgtca | 4380 |
| agtcagcgta | atgctctgcc | agtgttacaa | ccaattaacc | aattctgatt | agaaaaactc | 4440 |
| atcgagcatc | aaatgaaact | gcaatttatt | catatcagga | ttatcaatac | catatttttg | 4500 |
| aaaaagccgt | ttctgtaatg | aaggagaaaa | ctcaccgagg | cagttccata | ggatggcaag | 4560 |
| atcctggtat | cggtctgcga | ttccgactcg | tccaacatca | atacaaccta | ttaatttccc | 4620 |
| ctcgtcaaaa | ataaggttat | caagtgagaa | atcaccatga | gtgacgactg | aatccggtga | 4680 |
| gaatggcaaa | agcttatgca | tttctttcca | gacttgttca | acaggccagc | cattacgctc | 4740 |
| gtcatcaaaa | tcactcgcat | caaccaaacc | gttattcatt | cgtgattgcg | cctgagcgag | 4800 |
| acgaaatacg | cgatcgctgt | taaaaggaca | attacaaaca | ggaatcgaat | gcaaccggcg | 4860 |
| caggaacact | gccagcgcat | caacaatatt | ttcacctgaa | tcaggatatt | cttctaatac | 4920 |
|            |            |            |            |            |            |      |

```
ctggaatgct gttttcccgg ggatcgcagt ggtgagtaac catgcatcat caggagtacg 4980 gataaaatgc ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat 5040 ctcatctgta acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc 5100 atcgggcttc ccatacaatc gatagattgt cgcacctgat tgcccgacat tatcgcggc 5160 ccatttatac ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga 5220 cgtttcccgt tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag 5280 ttttattgtt catgatgata tatttttatc ttgtgcaatg tacacacaga gattttgaga 5340 cacaacgtgg cttccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 5400 gagcggatac atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt 5460 tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa 5520 aaataggcgt atcacagagc cctttcgtc 5549
```

<210> 82
<211> 5549
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 82

tegegegtht eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggeeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcq qtqtqaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960 taqaaqacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgacg taagtacege etatagacte tataggeaca eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catqqqctct tctccqqtaq cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga ccaccatggg 1920

|            |            |            |            |            | gcatgcgccg |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | agcacggcgc |      |
|            |            |            |            |            | aggcccagga |      |
|            |            |            |            |            | tgacctacaa |      |
|            |            |            |            |            | gcctgatcca |      |
|            |            |            |            |            | gctacttccc |      |
|            |            |            |            |            | tcggctggtg |      |
|            |            |            |            |            | gcgagaacac |      |
|            |            |            |            |            | aggtgctgga |      |
|            |            |            |            |            | accccgagta |      |
|            |            |            |            |            | catctgttgt |      |
|            |            |            |            |            | tcctttccta |      |
|            |            |            |            |            | tggggggtgg |      |
|            |            |            |            |            | ctggggatgc |      |
| ggtgggctct | atgggtaccc | aggtgctgaa | gaattgaccc | ggttcctcct | gggccagaaa | 2820 |
| gaagcaggca | catccccttc | tctgtgacac | accctgtcca | cgcccctggt | tcttagttcc | 2880 |
| agccccactc | ataggacact | catagctcag | gagggctccg | ccttcaatcc | cacccgctaa | 2940 |
| agtacttgga | gcggtctctc | cctccctcat | cagcccacca | aaccaaacct | agcctccaag | 3000 |
| agtgggaaga | aattaaagca | agataggcta | ttaagtgcag | agggagagaa | aatgcctcca | 3060 |
| acatgtgagg | aagtaatgag | agaaatcata | gaattttaag | gccatcatgg | ccttaatctt | 3120 |
| ccgcttcctc | gctcactgac | tegetgeget | cggtcgttcg | gctgcggcga | gcggtatcag | 3180 |
| ctcactcaaa | ggcggtaata | cggttatcca | cagaatcagg | ggataacgca | ggaaagaaca | 3240 |
| tgtgagcaaa | aggccagcaa | aaggccagga | accgtaaaaa | ggccgcgttg | ctggcgtttt | 3300 |
| tccataggct | cegeceecet | gacgagcatc | acaaaaatcg | acgctcaagt | cagaggtggc | 3360 |
| gaaacccgac | aggactataa | agataccagg | cgtttccccc | tggaagctcc | ctcgtgcgct | 3420 |
| ctcctgttcc | gaccctgccg | cttaccggat | acctgtccgc | ctttctccct | tcgggaagcg | 3480 |
| tggcgctttc | tcatagctca | cgctgtaggt | atctcagttc | ggtgtaggtc | gttcgctcca | 3540 |
| agctgggctg | tgtgcacgaa | cccccgttc  | agcccgaccg | ctgcgcctta | tccggtaact | 3600 |
| atcgtcttga | gtccaacccg | gtaagacacg | acttatcgcc | actggcagca | gccactggta | 3660 |
| acaggattag | cagagcgagg | tatgtaggcg | gtgctacaga | gttcttgaag | tggtggccta | 3720 |
| actacggcta | cactagaaga | acagtatttg | gtatctgcgc | tctgctgaag | ccagttacct | 3780 |
| tcggaaaaag | agttggtagc | tcttgatccg | gcaaacaaac | caccgctggt | agcggtggtt | 3840 |
| tttttgtttg | caagcagcag | attacgcgca | gaaaaaaagg | atctcaagaa | gatcctttga | 3900 |
| tcttttctac | ggggtctgac | gctcagtgga | acgaaaactc | acgttaaggg | attttggtca | 3960 |
| tgagattatc | aaaaaggatc | ttcacctaga | tccttttaaa | ttaaaaatga | agttttaaat | 4020 |
| caatctaaag | tatatatgag | taaacttggt | ctgacagtta | ccaatgctta | atcagtgagg | 4080 |
| cacctatctc | agcgatctgt | ctatttcgtt | catccatagt | tgcctgactc | 9999999999 | 4140 |
| ggcgctgagg | tctgcctcgt | gaagaaggtg | ttgctgactc | ataccaggcc | tgaatcgccc | 4200 |
| catcatccag | ccagaaagtg | agggagccac | ggttgatgag | agctttgttg | taggtggacc | 4260 |
| agttggtgat | tttgaacttt | tgctttgcca | cggaacggtc | tgcgttgtcg | ggaagatgcg | 4320 |
| tgatctgatc | cttcaactca | gcaaaagttc | gatttattca | acaaagccgc | cgtcccgtca | 4380 |
|            |            |            |            |            | agaaaaactc |      |
| atcgagcatc | aaatgaaact | gcaatttatt | catatcagga | ttatcaatac | catatttttg | 4500 |
| aaaaagccgt | ttctgtaatg | aaggagaaaa | ctcaccgagg | cagttccata | ggatggcaag | 4560 |
| atcctggtat | cggtctgcga | ttccgactcg | tccaacatca | atacaaccta | ttaatttccc | 4620 |
| ctcgtcaaaa | ataaggttat | caagtgagaa | atcaccatga | gtgacgactg | aatccggtga | 4680 |
| gaatggcaaa | agcttatgca | tttctttcca | gacttgttca | acaggccagc | cattacgctc | 4740 |
| gtcatcaaaa | tcactcgcat | caaccaaacc | gttattcatt | cgtgattgcg | cctgagcgag | 4800 |
|            |            |            |            |            | gcaaccggcg |      |
|            |            |            |            |            | cttctaatac |      |
| ctggaatgct | gttttcccgg | ggatcgcagt | ggtgagtaac | catgcatcat | caggagtacg | 4980 |
| gataaaatgc | ttgatggtcg | gaagaggcat | aaattccgtc | agccagttta | gtctgaccat | 5040 |
| ctcatctgta | acatcattgg | caacgctacc | tttgccatgt | ttcagaaaca | actctggcgc | 5100 |
| atcgggcttc | ccatacaatc | gatagattgt | cgcacctgat | tgcccgacat | tatcgcgagc | 5160 |
|            |            |            |            |            | tcgagcaaga |      |
|            |            |            |            |            |            |      |

```
cgtttcccgt tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag 5280
ttttattgtt catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga 5340
cacaacgtgg ctttcccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 5400
gagcggatac atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt 5460
tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa 5520
aaataggegt atcacgagge cetttegte
                                                                  5549
<210> 83
<211> 5549
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 83
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
tiggcagtac atclacgtat tagtcatcgc tattaccatg gigatgcggt titiggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaceteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080
tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtqcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga ccaccatggg 1920
cggcaagtgg agcaagagca gcgtgatcgg ctggcccgcc gtgcgcgagc gcatgcgccg 1980
```

egecgagece geogecgaeg gegtgggege egtgageege gaeetggaga ageaeggege 2040 eateaceage ageaeaceg eegecaacaa egecgeetge geetggetgg aggeeeagga 2100 ggaggaggag gtgggetteg eagtgaeege acaggtggea etgegegaa tgaeetacaa 2160 ggeegeegtg gaeetgagee aetteetgaa ggagaaggge ggeetggagg geetgateea 2220

|            |            | tcctggacct |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            | ccggccccgg |            |            |            |      |
|            |            | agcccgacaa |            |            |            |      |
|            |            | gcctgcacgg |            |            |            |      |
|            |            | tggccttcca |            |            |            |      |
|            |            | ccagatctgc |            |            |            |      |
|            |            | ccttgaccct |            |            |            |      |
|            |            | cgcattgtct |            |            |            |      |
|            |            | gggaggattg |            |            |            |      |
|            |            | aggtgctgaa |            |            |            |      |
|            |            | tctgtgacac |            |            |            |      |
|            |            | catagctcag |            |            |            |      |
|            |            | cctccctcat |            |            |            |      |
|            |            | agataggcta |            |            |            |      |
|            |            | agaaatcata |            |            |            |      |
|            |            | tcgctgcgct |            |            |            |      |
|            |            | cggttatcca |            |            |            |      |
|            |            | aaggccagga |            |            |            |      |
|            |            | gacgagcatc |            |            |            |      |
|            |            | agataccagg |            |            |            |      |
|            |            | cttaccggat |            |            |            |      |
|            |            | cgctgtaggt |            |            |            |      |
|            |            | cccccgttc  |            |            |            |      |
|            |            | gtaagacacg |            |            |            |      |
|            |            | tatgtaggcg |            |            |            |      |
|            |            | acagtatttg |            |            |            |      |
|            |            | tcttgatccg |            |            |            |      |
|            |            | attacgcgca |            |            |            |      |
|            |            | gctcagtgga |            |            |            |      |
|            |            | ttcacctaga |            |            |            |      |
|            |            | taaacttggt |            |            |            |      |
|            |            | ctatttcgtt |            |            |            |      |
|            |            | gaagaaggtg |            |            |            |      |
|            |            | agggagccac |            |            |            |      |
|            |            | tgctttgcca |            |            |            |      |
|            |            | gcaaaagttc |            |            |            |      |
|            |            | agtgttacaa |            |            |            |      |
|            |            | gcaatttatt |            |            |            |      |
|            |            | aaggagaaaa |            |            |            |      |
|            |            | ttccgactcg |            |            |            |      |
|            |            | caagtgagaa |            |            |            |      |
|            |            | tttctttcca |            |            |            |      |
|            |            | caaccaaacc |            |            |            |      |
|            |            | taaaaggaca |            |            |            |      |
|            |            | caacaatatt |            |            |            |      |
| ctggaatgct | gttttcccgg | ggatcgcagt | ggtgagtaac | catgcatcat | caggagtacg | 4980 |
| gataaaatgc | ttgatggtcg | gaagaggcat | aaattccgtc | agccagttta | gtctgaccat | 5040 |
|            |            | caacgctacc |            |            |            |      |
| accgggcttc | ccaracaarc | gatagattgt | cgcacctgat | tgcccgacat | tatcgcgagc | 5160 |
| ccacccacac | ccatataaat | cagcatccat | gccggaactt | aatcgcggcc | ccgagcaaga | 5220 |
| egetececgt | rgaarargge | tcataacacc | ccttgtatta | ccgtttatgt | aagcagacag | 5280 |
| ccccattgtt | catgatgata | tatttttatc | ctgtgcaatg | caacatcaga | gattttgaga | 5340 |
| cacaacgtgg | GEEECCCCC  | cccccatta  | ccgaagcatt | tatcagggtt | attgtctcat | 5400 |
| gageggatae | acacccgaac | gtatttagaa | aaataaacaa | acaggggttc | cgcgcacatt | 5460 |
| rccccgaaaa | grgccaccrg | acgtctaaga | aaccattatt | accatgacat | taacctataa | 5520 |

aaataggcgt atcacgaggc cctttcgtc

<210> 84

5549

```
<211> 6609
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegttt eggtgatgae ggtgaaaaee tetgaeaeat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccqcqtt acataactta cqqtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
cataqtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacqtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcqqt tttgqcaqta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattqa 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcqtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tettatgeat getatactgt tittggettg gggeetatac accecegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct totocggtag cggcggagct tocacatccg agccctggtc ccatgcctcc 1500
ageggeteat ggtegetegg eageteettg eteetaacag tggaggeeag acttaggeac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgage gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt teetttecat gggtetttte 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt gaaggagaag taccagcacc tgtggcgctg gggctggcgc tqqqgcacca 1980
tgctgctggg catgctgatg atctgcagcg ccaccgagaa gctgtgggtg accqtgtact 2040
acggcgtgcc cgtgtggcgc gaggccacca ccaccctgtt ctgcgccagc gacgccaagg 2100
cetacgacae egaggtgeae aacgtgtggg ceaeceaege etgegtgeee accqaeecea 2160
acceccagga ggtggtgctg ggcaacgtga ccgagaactt caacatgtgg aagaacaaca 2220
tggtggacca gatgcacgag gacatcatca gcctgtggga cgagagcctg aagccctgcg 2280
tgaagctgac ccccctgtgc gtgaccctga actgcaccaa cctgaacatc accaagaaca 2340
ccaccaacct gaccagcagc agctggggca tgatggagga gggcgagatc aagaactgca 2400
gcttctacat caccaccage atccgcaaca aggtgaagaa ggagtacqcc ctqttcaacc 2460
gcctggacgt ggtgcccgtg aagaacacca gcaacaccaa gtaccqcctq atcaqctqca 2520
```

- 196 -

| acaccacat  | astasaaaaa | acabaaaaa  |            |            |             |              |
|------------|------------|------------|------------|------------|-------------|--------------|
| acaccagege | gatcacccag | geeegeeea  | aggrgagett | ccagcccatc | cccatccact  | 2580         |
| actgegtgee | cgccggcttc | gccatcctga | agtgcaacaa | caagaccttc | aacggcagcg  | 2640         |
| geceetgeae | caacgtgagc | accgtgcagt | gcacccacgg | cateegeece | gtggtgagca  | 2700         |
| cccagctgct | gctgaacggc | agcctggccg | aggaggacat | cgtgatccgc | agcgaggact  | 2760         |
| tcaccgacaa | cgtgaagacc | atcatcgtgc | agctgaacga | gagcgtggtg | atcaactgca  | 2820         |
| cccgccccaa | caacaacacc | cgcgagcgcc | tgagcatcgg | ccccggccgc | gccttctacg  | 2880         |
| cccgccgcaa | catcatcggc | gacatecgee | aggcccactg | caacatcagc | cgcgccaagt  | 2940         |
| ggaacaacac | cctgcagcag | atcgtgatca | agctgcgcga | gaagttccgc | aacaagacca  | 3000         |
| tegeetteaa | ccagagcagc | ggcggcgacc | ccgagatcgt | gatgcacagc | ttcaactgcg  | 3060         |
| gcggcgagtt | cttctactgc | aacaccgccc | agctgttcaa | cagcacctgg | aacgtggccg  | 3120         |
| gcggcaccaa | cggcaccgag | ggcaacgaca | tcatcaccct | gcagtgccgc | atcaagcaga  | 3180         |
| tcatcaacat | gtggcagaag | gtgggcaagg | ccatgtacgc | ccccccatc  | accggccaga  | 3240         |
| tccgctgcag | cagcaacatc | accggcctgc | tgctgacccg | cgacggcggc | aacagcaccg  | 3300         |
| agaccgagac | cgagatcttc | cgccccggcg | gcggcgacat | gcgcgacaac | tggcgcagcg  | 3360         |
| agctgtacaa | gtacaaggtg | gtgcgcatcg | agcccatcgg | cgtggccccc | acccgcgcca  | 3420         |
| agcttaccgt | gcaggcccgc | ctgctgctga | gcggcatcgt | gcagcagcag | aacaacctgc  | 3480         |
| tgcgcgccat | cgaggcccag | cagaacatgc | tgcgcctgac | cgtgtggggc | atcaagcagc  | 3540         |
| tgcaggcccg | cgtgctggcc | ctggagcgct | acctgcgcga | ccagcagctc | gagtgaggat  | 3600         |
| ccagatctgc | tgtgccttct | agttgccagc | catctgttgt | ttgcccctcc | cccqtqcctt  | 3660         |
| ccttgaccct | ggaaggtgcc | actcccactg | tcctttccta | ataaaatgag | gaaattqcat  | 3720         |
| cgcattgtct | gagtaggtgt | cattctattc | tggggggtgg | ggtggggcag | gacagcaagg  | 3780         |
| gggaggattg | ggaagacaat | agcaggcatg | ctggggatgc | ggtgggctct | atgggtaccc  | 3840         |
| aggtgctgaa | gaattgaccc | ggttcctcct | gggccagaaa | gaagcaggca | catccccttc  | 3900         |
| tctgtgacac | accctgtcca | cgcccctggt | tettagttee | agccccactc | ataggacact  | 3960         |
| catageteag | gagggctccg | ccttcaatcc | cacccgctaa | agtacttgga | gcggtctctc  | 4020         |
| cctccctcat | cagcccacca | aaccaaacct | agectecaag | agtgggaaga | aattaaagca  | 4080         |
| agataggcta | ttaagtgcag | agggagagaa | aatgcctcca | acatgtgagg | aagtaatgag  | 4140         |
| agaaatcata | gaattttaag | gccatcatgg | ccttaatctt | ccacttcctc | gctcactgac  | 4200         |
| tegetgeget | cggtcgttcg | gctgcggcga | gcggtatcag | ctcactcaaa | gacaataata  | 4260         |
| cggttatcca | cagaatcagg | ggataacgca | qqaaaqaaca | totoagcaaa | aggccagcaa  | 4320         |
| aaggccagga | accgtaaaaa | ggccgcgttg | ctagcatttt | tccataggct | cogocococt  | 4380         |
| gacgagcatc | acaaaaatcg | acqctcaaqt | cagaggtggc | gaaacccgac | aggachahaa  | 4440         |
| agataccagg | cgtttccccc | tggaagetee | ctcatacact | ctcctattcc | gaccctgccg  | 4500         |
| cttaccqqat | acctgtccgc | ctttctccct | tcagaaaaca | tageactite | tcatagetca  | 4560         |
| cgctqtaggt | atctcagttc | ggtgtaggtc | gttcgctcca | agctgggctg | tatacacaaa  | 4620         |
| cccccqttc  | agcccgaccg | ctgcgcctta | tccggtaact | atcatcttca | atccaaccca  | 4680         |
| qtaaqacacq | acttatcgcc | actogcagca | accactaata | acaggattag | Georgage    | 4740         |
| tatgtaggeg | gtgctacaga | attettaaaa | taataaccta | actacaacta | cactagagg   | 4900         |
| acagtatttg | gtatctgcgc | tctgctgaag | ccaottacct | tcggaaaaa  | agttagtage  | 4960         |
| tettgatecg | gcaaacaaac | caccactant | agcggtggtt | ttttatta   | caaccagcagc | 4020         |
| attacococa | gaaaaaaagg | atctcaagaa | gateetttga | tettttetae | aggatataa   | 4920         |
| actcaataaa | acgaaaactc | acattaagaa | artttootca | tgagattatg | agageeegae  | 4300<br>E040 |
| ttcacctaga | tccttttaaa | ttaaaaatga | agttttaaat | caatchaaag | tabababasa  | 5100         |
| taaacttogt | ctgacagtta | ccaatactta | atcacteace | caacccaaag | acatatyay   | 2100         |
| chatttcgtt | catccatagt | tacctaecta | accagegagg | cacccaccc  | agegatetgt  | 2700         |
| ceacecegee | ttgctgactc | ataccagge  | 9999999999 | ggcgctgagg | cctgcctcgt  | 5220         |
| adddaddcac | aattaataaa | acaccaggee | tagategace | cateatecag | ccagaaagtg  | 5280         |
| tactttacca | ggttgatgag | tagattatag | caggiggace | agetggtgat | cccgaacccc  | 5340         |
| gcasaachto | cggaacggtc | agaaagagag | ggaagacgcg | cgacctgatc | ccccaaccca  | 5400         |
| antottacaa | gatttattca | acadayccyc | -greeegrea | agtcagcgta | argetetgee  | 5460         |
| gcaabbtath | ccaattaacc | thata      | ayaaaaaccc | accgagcatc | aaatgaaact  | 5520         |
| guadutatt  | catatcagga | coattacet  | catacttttg | aaaaagccgt | ttctgtaatg  | 5580         |
| ttaaaaata  | ctcaccgagg | cagetecata | ggatggcaag | accctggtat | cggtctgcga  | 5640         |
| anach      | tccaacatca | acacaaccca | traatttccc | ctcgtcaaaa | acaaggttat  | 5700         |
| tttattt    | atcaccatga | ycgacgactg | aacccggtga | gaatggcaaa | agcttatgca  | 5760         |
| LLLCLEECCA | gacttgttca | acaggccagc | cattacgete | gtcatcaaaa | tcactcgcat  | 5820         |

```
caaccaaacc gttattcatt cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt 5880 taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact gccagcgcat 5940 caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct gttttcccgg 6000 ggatcgcagt ggtgagtaac catgcatcat caggagtacg gataaaaatgc ttgatggtcg 6120 caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc ccatacaatc 6180 gatagattgt cgcacctgat tgcccgacat tatcgcggcc ccatttatac ccatataaat 6240 cagcatccat gttggaattt actgttatgt aaccacaagcacg ccttgatta ctgttatgt aacaccaga cgtttcccgt tgaatatggc 6300 tattttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg ctttccccc 6420 cccccatta ttgaagcatt tatcagggtt attgtccat gagcggatac atatttgaat 6480 gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa gtgccacctg 6540 acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc 6600 cctttcgtc
```

<210> 85 <211> 6783 <212> DNA <213> Artificial Sequence <223> plasmid pVR1012x/s containing HIV genes togogogttt oggtgatgac ggtgaaaacc totgacacat gcagetcceg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataaqcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetqtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattq gaacgcqqat 1020 teccegtgee aagagtgacg taagtacege ctatagacte tataggeaca eccetttgge 1080 tottatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

| attacaatac  | tattaacaat | aasaaacsat  | ataatataaa | cagtactcgt | tactaccaca | 1900 |
|-------------|------------|-------------|------------|------------|------------|------|
| CGCGCCACCA  | gacataatag | ctoacagact  | aacadactdt | tcctttccat | agatettta  | 1060 |
| tocautcacc  | gtcataatag | catatastas  | aacagaccgc | ccccccac   | gggtetttte | 1000 |
| ccatacacat  | geegeegaea | tagaggacca  | bataccaca  | ccgctctaga | gatategeea | 1920 |
| tactactaca  | gaaggagaag | caccagcacc  | cgcggcgccg | gggctggcgc | rggggcacca | 1980 |
| rgergerggg  | catgetgatg | acetgeageg  | ccaccgagaa | gctgtgggtg | accgtgtact | 2040 |
| acggegegee  | cgrgrggcgc | gaggccacca  | ccaccctgtt | ctgcgccagc | gacgccaagg | 2100 |
| cctacgacac  | cgaggtgcac | aacgtgtggg  | ccacccacgc | ctgcgtgccc | accgacccca | 2160 |
| acccccagga  | ggtggtgctg | ggcaacgtga  | ccgagaactt | caacatgtgg | aagaacaaca | 2220 |
| tggtggacca  | gatgcacgag | gacatcatca  | gcctgtggga | cgagagcctg | aagccctgcg | 2280 |
| tgaagctgac  | cccctgtgc  | gtgaccctga  | actgcaccaa | cctgaacatc | accaagaaca | 2340 |
| ccaccaacct  | gaccagcagc | agctggggca  | tgatggagga | gggcgagatc | aagaactgca | 2400 |
| gcttctacat  | caccaccagc | atccgcaaca  | aggtgaagaa | ggagtacgcc | ctgttcaacc | 2460 |
| gcctggacgt  | ggtgcccgtg | aagaacacca  | gcaacaccaa | gtaccgcctg | atcagctgca | 2520 |
| acaccagcgt  | gatcacccag | gcctgcccca  | aggtgagctt | ccagcccatc | cccatccact | 2580 |
| actgcgtgcc  | cgccggcttc | gccatcctga  | agtgcaacaa | caagaccttc | aacggcagcg | 2640 |
| gcccctgcac  | caacgtgagc | accgtgcagt  | gcacccacgg | catccgcccc | gtggtgagca | 2700 |
| cccagctgct  | gctgaacggc | agcctggccg  | aggaggacat | cgtgatccgc | agcgaggact | 2760 |
| tcaccgacaa  | cgtgaagacc | atcatcgtgc  | agctgaacga | gagcgtggtg | atcaactgca | 2820 |
| cccgccccaa  | caacaacacc | cgcgagcgcc  | tgagcatcgg | ccccggccgc | gccttctacq | 2880 |
| cccqccqcaa  | catcatcggc | gacatccgcc  | aggcccactg | caacatcagc | cacaccaaat | 2940 |
| ggaacaacac  | cctgcagcag | atcataatca  | agctgcgcga | gaagttccgc | aacaagacca | 3000 |
| tcqccttcaa  | ccagagcagc | ggcggcgacc  | ccgagatcgt | gatgcacagc | ttcaactgcg | 3060 |
| gcggcgagtt  | cttctactgc | aacaccgccc  | agctgttcaa | cagcacctgg | aacataacca | 3120 |
| acaacaccaa  | caacaccaaa | adcaacdaca  | tcatcaccct | gcagtgccgc | atcaaccaca | 3190 |
| tcatcaacat  | ataacaaaaa | ataaacsaaa  | ccatchaccc | ccccccatc  | accaagcaga | 2700 |
| tecactacaa  | caccaacate | acconcetoe  | tactascaa  | cgacggcggc | accegecaga | 3240 |
| adaccdadac  | cascatette | caccacaca   | geograpast | acacasassa | tagggagg   | 3300 |
| agacegagae  | atacaaaata | atacacataa  | geggegaeae | gcgcgacaac | rggegeageg | 3360 |
| agettaceat  | gcacaaggcg | gracegearca | ageceategg | cgtggccccc | accegegeea | 3420 |
| tacacacac   | gcaggcccgc | cogocococo  | baggeategt | gcagcagcag | aacaacctgc | 3480 |
| tacagagagag | cgaggcccag | cagaacacge  | tgegeetgae | cgtgtggggc | accaagcagc | 3540 |
| cgcaggeeeg  | egrgerggee | ctggageget  | acctgcgcga | ccagcagctc | gaggacatct | 3600 |
| ggaacaacat  | gacciggatg | gagugggage  | gcgagaccga | caactacacc | gactacatct | 3660 |
| acgacctgct  | ggagaagagc | cagacccagc  | aggagaagaa | cgagaaggag | ctgctggagc | 3720 |
| tggacaagtg  | ggccagcctg | tggaactggt  | tcgacatcac | caactggtga | ggatccagat | 3780 |
| ctgctgtgcc  | ttctagttgc | cagccatctg  | ttgtttgccc | ctccccgtg  | ccttccttga | 3840 |
| ccctggaagg  | tgccactccc | actgtccttt  | cctaataaaa | tgaggaaatt | gcatcgcatt | 3900 |
| gtctgagtag  | gtgtcattct | attctggggg  | gtggggtggg | gcaggacagc | aagggggagg | 3960 |
| attgggaaga  | caatagcagg | catgctgggg  | atgcggtggg | ctctatgggt | acccaggtgc | 4020 |
| tgaagaattg  | acccggttcc | tcctgggcca  | gaaagaagca | ggcacatccc | cttctctgtg | 4080 |
| acacaccctg  | tccacgcccc | tggttcttag  | ttccagcccc | actcatagga | cactcatage | 4140 |
| tcaggagggc  | tccgccttca | atcccacccg  | ctaaagtact | tggagcggtc | tctccctccc | 4200 |
| tcatcagccc  | accaaaccaa | acctagcctc  | caagagtggg | aagaaattaa | agcaagatag | 4260 |
| gctattaagt  | gcagaggag  | agaaaatgcc  | tccaacatgt | gaggaagtaa | tgagagaaat | 4320 |
| catagaattt  | taaggccatc | atggccttaa  | tcttccgctt | cctcgctcac | tgactcgctg | 4380 |
| cgctcggtcg  | ttcggctgcg | gcgagcggta  | tcagctcact | caaaggcggt | aatacggtta | 4440 |
| tccacagaat  | caggggataa | cgcaggaaag  | aacatgtgag | caaaaggcca | gcaaaaggcc | 4500 |
| aggaaccgta  | aaaaggccgc | gttgctggcg  | tttttccata | ggctccgccc | ccctgacgag | 4560 |
| catcacaaaa  | atcgacgctc | aagtcagagg  | tggcgaaacc | cgacaggact | ataaagatac | 4620 |
| caggcgtttc  | cccctggaag | ctccctcgtg  | cgctctcctg | ttccgaccct | gccgcttacc | 4680 |
| ggatacctgt  | ccgcctttct | cccttcggga  | agegtggege | tttctcatag | ctcacactat | 4740 |
| aggtatctca  | gttcggtgta | ggtcgttcac  | tccaaqctqq | gctgtgtgca | cgaacccccc | 4800 |
| gttcagcccq  | accqctqcqc | cttatccoot  | aactatcotc | ttgagtccaa | cccaataaaa | 4860 |
| cacgacttat  | cgccactage | agcagccact  | ggtaacagga | ttagcagage | gaggtatota | 4920 |
| ggcggtgcta  | cagagttett | gaagtggtgg  | cctaactaco | gctacactag | aagaacadta | 4980 |
| tttggtatct  | gegetetach | gaagccaght  | accttcooas | aaagagttgg | tagetettes | 5040 |
|             | 5.505      |             |            | ~~~3~3~5   | gullinga   | 2040 |

```
teeggeaaac aaaccacege tggtageggt ggtttttttg tttgeaagea geagattaeg 5100
cgcagaaaaa aaggatetea agaagateet ttgatetttt etaeggggte tgaegeteag 5160
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaaag gatcttcacc 5220
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 5280
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 5340
cgttcatcca tagttgcctg actcgggggg ggggggcgct gaggtctgcc tcgtgaagaa 5400
ggtgttgctg actcatacca ggcctgaatc gccccatcat ccagccagaa agtgagggag 5460
ccacggttga tgagagcttt gttgtaggtg gaccagttgg tgattttgaa cttttgcttt 5520
gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct gatccttcaa ctcagcaaaa 5580
gttcgattta ttcaacaaag ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt 5640
acaaccaatt aaccaattct gattagaaaa actcatcgag catcaaatga aactgcaatt 5700
tattcatatc aggattatca ataccatatt tttqaaaaag ccgtttctqt aatgaaggag 5760
aaaactcacc gaggcagttc cataggatgg caagatcctg gtatcggtct gcgattccqa 5820
ctcgtccaac atcaatacaa cctattaatt tcccctcgtc aaaaataagg ttatcaagtg 5880
agaaatcacc atgagtgacg actgaatccg gtgagaatgg caaaagctta tgcatttctt 5940
tecagacttg tteaacagge cagecattae getegteate aaaateacte geateaacea 6000
aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaag 6060
gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa 6120
tattttcacc tgaatcagga tattcttcta atacctggaa tgctgttttc ccggqqatcq 6180
cagtggtgag taaccatgca tcatcaggag tacggataaa atgcttgatg gtcggaagag 6240
qcataaattc cqtcaqccaq tttaqtctqa ccatctcatc tqtaacatca ttqqcaacqc 6300
tacctttgcc atgtttcaga aacaactctg gcgcatcggg cttcccatac aatcgataga 6360
ttgtcgcacc tgattgcccg acattatcgc gagcccattt atacccatat aaatcagcat 6420
ccatgttgga atttaatcgc ggcctcgagc aagacgtttc ccgttgaata tggctcataa 6480
caccccttgt attactgttt atgtaagcag acagttttat tgttcatgat gatatatttt 6540
tatettgtge aatgtaacat cagagatttt gagacacaac gtggetttee ceeeecece 6600
attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 6660
agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacqtct 6720
aagaaaccat tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc 6780
gtc
```

```
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgeg gtgtgaaata eegeadagat gegtaaggag aaaatacege atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatqq 420
cccqcctqqc tqaccqccca acgacccccq cccattqacq tcaataatqa cqtatqttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacqtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
```

. <210> 86
 <211> 6861

|            |                |            |            |            |            | 040  |
|------------|----------------|------------|------------|------------|------------|------|
| cgtcaatggg | agtttgtttt     | ggcaccaaaa | ccaacgggac | tttccaaaat | gtcgtaacaa | 840  |
| ctccgcccca | ttgacgcaaa     | tgggcggtag | gcgtgtacgg | tgggaggtct | atataagcag | 900  |
| agctcgttta | gtgaaccgtc     | agatcgcctg | gagacgccat | ccacgctgtt | ttgacctcca | 960  |
|            |                |            |            |            | gaacgcggat |      |
| tccccgtgcc | aagagtgacg     | taagtaccgc | ctatagactc | tataggcaca | cccctttggc | 1080 |
| tcttatgcat | gctatactgt     | ttttggcttg | gggcctatac | acccccgctt | ccttatgcta | 1140 |
|            |                |            |            |            | accactcccc |      |
| tattggtgac | gatactttcc     | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260 |
| tattggctat | atgccaatac     | tetgteette | agagactgac | acggactctg | tatttttaca | 1320 |
| ggatggggtc | ccatttatta     | tttacaaatt | cacatataca | acaacgccgt | cccccgtgcc | 1380 |
| cgcagttttt | attaaacata     | gcgtgggatc | tccacgcgaa | tctcgggtac | gtgttccgga | 1440 |
|            |                |            |            |            | ccatgcctcc |      |
| agcggctcat | ggtcgctcgg     | cagctccttg | ctcctaacag | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac     | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg     | ggctcgcacg | gctgacgcag | atggaagact | taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag     | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt     | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg | 1800 |
| cgcgccacca | gacataatag     | ctgacagact | aacagactgt | tectttecat | gggtcttttc | 1860 |
| tgcagtcacc | gtcgtcgaca     | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatcgcca | 1920 |
| ccatgcgcgt | gaaggagaag     | taccagcacc | tgtggcgctg | gggctggcgc | tggggcacca | 1980 |
| tgctgctggg | catgctgatg     | atctgcagcg | ccaccgagaa | gctgtgggtg | accgtgtact | 2040 |
| acggcgtgcc | cgtgtggcgc     | gaggccacca | ccaccctgtt | ctgcgccagc | gacgccaagg | 2100 |
| cctacgacac | cgaggtgcac     | aacgtgtggg | ccacccacgc | ctgcgtgccc | accgacccca | 2160 |
| acccccagga | ggtggtgctg     | ggcaacgtga | ccgagaactt | caacatgtgg | aagaacaaca | 2220 |
| tggtggacca | qatqcacgag     | gacatcatca | gcctgtggga | cgagagcctg | aagccctgcg | 2280 |
| tgaagctgac | cccctgtgc      | gtgaccctga | actgcaccaa | cctgaacatc | accaagaaca | 2340 |
| ccaccaacct | gaccagcagc     | agctggggca | tgatggagga | gggcgagatc | aagaactgca | 2400 |
| gcttctacat | caccaccagc     | atccgcaaca | aggtgaagaa | ggagtacgcc | ctgttcaacc | 2460 |
| geetggaegt | gatacccata     | aagaacacca | gcaacaccaa | gtaccgcctg | atcagctgca | 2520 |
| acaccagcgt | gatcacccag     | gcctgcccca | aggtgagctt | ccagcccatc | cccatccact | 2580 |
| actocotoco | caccaacttc     | gccatcctga | agtgcaacaa | caagaccttc | aacggcagcg | 2640 |
| gcccctgcac | caacqtqaqc     | accetecagt | gcacccacgg | catecgeece | gtggtgagca | 2700 |
| cccagetget | actgaacggc     | agectggeeg | aggaggacat | cgtgatccgc | agcgaggact | 2760 |
| tcaccgacaa | cqtqaaqacc     | atcatcgtgc | agctgaacga | gagcgtggtg | atcaactgca | 2820 |
| ccccccaa   | caacaacacc     | cqcqaqcqcc | tgagcatcgg | ccccggccgc | gccttctacg | 2880 |
| cccqccqcaa | catcatcggc     | gacatccgcc | aggcccactg | caacatcagc | cgcgccaagt | 2940 |
| ggaacaacac | cctqcaqcaq     | atcgtgatca | agctgcgcga | gaagttccgc | aacaagacca | 3000 |
|            |                |            |            |            | ttcaactgcg |      |
| acaacaaatt | cttctactqc     | aacaccqccc | agctgttcaa | cagcacctgg | aacgtggccg | 3120 |
| gcggcaccaa | cqqcaccqaq     | ggcaacgaca | tcatcaccct | gcagtgccgc | atcaagcaga | 3180 |
| tcatcaacat | gtggcagaag     | gtgggcaagg | ccatgtacgc | ccccccatc  | accggccaga | 3240 |
| tecactacaa | cagcaacatc     | accggcctgc | tgctgacccg | cgacggcggc | aacagcaccg | 3300 |
| agaccgagac | cgagatette     | caccccaaca | gcggcgacat | gcgcgacaac | tggcgcagcg | 3360 |
| agctgtacaa | gtacaaggtg     | gtgcgcatcg | agcccatcgg | cgtggccccc | acccgcgcca | 3420 |
|            |                |            |            |            | aacaacctgc |      |
| tacacaccat | cgaggcccag     | cagaacatgo | tgcgcctgac | catatagaac | atcaagcagc | 3540 |
| tacagacca  | catactaacc     | ctggagcgct | acctgcgcga | ccaqcagctc | gaggacatct | 3600 |
| ggaacaacat | gacctggatg     | gagtagaage | gegagatega | caactacacc | gactacatct | 3660 |
| acgacctgct | ggagaagag      | cagacccagc | aggagaagaa | cqaqaaqqaq | ctgctggagc | 3720 |
| tagacaagto | gaccaaccta     | tagaactagt | tcgacatcac | caactqqctq | tggtacatcc | 3780 |
| acctattcat | catgatcgtg     | gacgacctaa | teggeetaca | catcatatto | gccgtgctga | 3840 |
| gcatctgagg | atccagatct     | gctgtgcctt | ctagttqcca | gccatctqtt | gtttgcccct | 3900 |
| coccataco  | ttccttgacc     | ctagaaggta | ccactcccac | tatectttee | taataaaatg | 3960 |
| aggaaattgc | atcocattot     | ctgagtaggt | gtcattctat | tctqqqqqat | ggggtggggc | 4020 |
| aggacagcaa | aaaaaaaaa      | taggaagaca | atagcaggca | tactagagat | gcggtgggct | 4080 |
| ~          | יייכביים ככככנ | -55555-    | J          |            | J - J J    | •    |

```
ctatgggtac ccaggtgctg aagaattgac ccggttcctc ctgggccaga aagaagcagg 4140
cacateceet tetetgtgae acaceetgte caegeeeetg gttettagtt ceageeeeac 4200
teataggaca eteatagete aggagggete egeetteaat eccaceeget aaagtaettg 4260
gageggtete tecetecete ateageceae caaaccaaac etageeteea agagtgggaa 4320
gaaattaaag caagataggc tattaagtgc agagggagag aaaatgcctc caacatgtga 4380
ggaagtaatg agagaaatca tagaatttta aggccatcat ggccttaatc ttccgcttcc 4440
tegeteactg actegetgeg eteggtegtt eggetgegge gageggtate ageteactea 4500
aaggeggtaa taeggttate cacagaatea ggggataaeg caggaaagaa catgtgagea 4560
aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg 4620
ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg 4680
acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 4740
ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 4800
teteataget caegetgtag gtateteagt teggtgtagg tegttegete caagetggge 4860
tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcqtctt 4920
gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 4980
agcagagega ggtatgtagg eggtgetaca gagttettga agtggtggee taactaegge 5040
tacactagaa gaacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa 5100
agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg tttttttgtt 5160
tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct 5220
acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta 5280
tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa 5340
agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga qqcacctatc 5400
tcagcgatct gtctatttcg ttcatccata gttgcctgac tcgggggggg ggggcgctga 5460
ggtctgcctc gtgaagaagg tgttgctgac tcataccagg cctgaatcgc cccatcatcc 5520
agccagaaag tgagggagcc acggttgatg agagctttgt tgtaggtgga ccagttggtg 5580
attttgaact tttgctttgc cacggaacgg tctgcgttgt cgggaagatg cgtgatctga 5640
teetteaact cagcaaaagt tegatttatt caacaaagee geegteeegt caagteageg 5700
taatgctctg ccagtgttac aaccaattaa ccaattctga ttagaaaaac tcatcgagca 5760
tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatttt tgaaaaagcc 5820
gtttctgtaa tgaaggagaa aactcaccga ggcagttcca taggatggca agatcctqqt 5880
ateggtetge gatteegact egteeaacat caatacaace tattaattte ecetegteaa 5940
aaataaggtt atcaagtgag aaatcaccat gagtgacgac tgaatccggt gagaatggca 6000
aaagettatg cathtettte cagaettgtt caacaggeca gecattacge tegteatcaa 6060
aatcactcgc atcaaccaaa ccgttattca ttcgtgattg cgcctgagcg agacgaaata 6120
cgcgatcgct gttaaaagga caattacaaa caggaatcga atgcaaccgg cgcaggaaca 6180
ctgccagcgc atcaacaata ttttcacctg aatcaggata ttcttctaat acctggaatg 6240
ctgttttccc ggggatcgca gtggtgagta accatgcatc atcaggagta cggataaaat 6300
gcttgatggt cggaagaggc ataaattccg tcagccagtt tagtctgacc atctcatctg 6360
taacatcatt ggcaacgcta cctttgccat gtttcagaaa caactctggc gcatcgggct 6420
teccatacaa tegatagatt gtegeacetg attgeecegae attategega geccatttat 6480
acccatataa atcagcatcc atgttggaat ttaatcgcgg cctcgagcaa gacgtttccc 6540
gttgaatatg gctcataaca ccccttgtat tactgtttat gtaagcagac agttttattg 6600
ttcatgatga tatattttta tcttgtgcaa tgtaacatca gagattttga gacacaacgt 6660
ggctttcccc cccccccat tattgaagca tttatcaggg ttattgtctc atgagcggat 6720
acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa 6780
aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc 6840
gtatcacgag gccctttcgt c
```

<21,0> 87

<211> 7552

<212> DNA

<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 87 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 ecegeetgge tgacegeeea aegaceeeeg eceattgaeg teaataatga egtatgttee 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecce 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae aeggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgage gtggagattg ggetegeacg getgaegeag atggaagaet taaggeageg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgcgt gaaggagaag taccagcacc tgtggcgctg gggctggcgc tggggcacca 1980 tgctgctggg catgctgatg atctgcagcg ccaccgagaa gctgtgggtg accgtgtact 2040 acggcgtgcc cgtgtggcgc gaggccacca ccaccctgtt ctgcgccagc gacgccaagg 2100 cctacgacac cgaggtgcac aacgtgtggg ccacccacgc ctgcgtgccc accgacccca 2160 acccccagga ggtggtgctg ggcaacgtga ccgagaactt caacatgtgg aagaacaaca 2220 tggtggacca gatgcacgag gacatcatca gcctgtggga cgagagcctg aagccctgcg 2280 tgaagctgac ccccctgtgc gtgaccctga actgcaccaa cctgaacatc accaagaaca 2340 ccaccaacct gaccagcagc agctggggca tgatggagga gggcgagatc aagaactgca 2400 gettetacat caccaccage atcegeaaca aggtgaagaa ggagtacgee etgtteaace 2460 gcctggacgt ggtgcccgtg aagaacacca gcaacaccaa gtaccgcctg atcagctgca 2520 acaccagegt gateacccag geetgeecca aggtgagett ccageccate eccatecact 2580 actgcgtgcc cgccggcttc gccatcctga agtgcaacaa caagaccttc aacggcagcg 2640 gcccctgcac caacgtgagc accgtgcagt gcacccacgg catccgcccc gtggtgagca 2700 eccagetget getgaaegge ageetggeeg aggaggaeat egtgateege agegaggaet 2760 tcaccgacaa cgtgaagacc atcatcgtgc agctgaacga gagcgtggtg atcaactgca 2820 cccgcccaa caacaacac cgcgagcgcc tgagcatcgg ccccggccgc gccttctacg 2880 cccgccgcaa catcatcggc gacatccgcc aggcccactg caacatcagc cgcgccaagt 2940 ggaacaacac cctgcagcag atcgtgatca agctgcgcga gaagttccgc aacaagacca 3000 tegeetteaa ecagageage ggeggegace ecgagategt gatgeacage tteaactgeg 3060

| gcggcgagtt | cttctactgc | aacaccgccc | agctgttcaa | cagcacctgg | aacgtggccg | 3120 |
|------------|------------|------------|------------|------------|------------|------|
|            | cggcaccgag |            |            |            |            |      |
|            | gtggcagaag |            |            |            |            |      |
|            | cagcaacatc |            |            |            |            |      |
|            | cgagatcttc |            |            |            |            |      |
|            | gtacaaggtg |            |            |            |            |      |
|            | cgtgcagcgc |            |            |            |            |      |
|            | cgccggcage |            |            |            |            |      |
|            | gagcggcatc |            |            |            |            |      |
|            | gctgcgcctg |            |            |            |            |      |
| ccctqqaqcq | ctacctgcgc | gaccagcagc | tgatgggcat | ctagaactac | agcggcaagc | 3720 |
|            | caccagegtg |            |            |            |            |      |
|            | catgacetgg |            |            |            |            |      |
|            | gctggagaag |            |            |            |            |      |
|            | gtgggccagc |            |            |            |            |      |
|            | catcatgate |            |            |            |            |      |
|            | gaaccgcgtg |            |            |            |            |      |
|            | cggccccgac |            |            |            |            |      |
|            | cggccccagc |            |            |            |            |      |
|            | gttcagctac |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | gggccgccgc |            |            |            |            |      |
|            | ggagctgaag |            |            |            |            |      |
|            | cgtgcaggcc |            |            |            |            |      |
|            | ggagaccctg |            |            |            |            |      |
|            | cctggagctg |            |            |            |            |      |
|            | agccatctgt |            |            |            |            |      |
| geeaeteeea | ctgtcctttc | Ctaataaaat | gaggaaattg | eategeattg | tetgagtagg | 4680 |
| cgccacccca | ttctgggggg | r999919999 | caggacagca | agggggagga | ttgggaagac | 4740 |
|            | atgctgggga |            |            |            |            |      |
|            | cctgggccag |            |            |            |            |      |
|            | ggttcttagt |            |            |            |            |      |
|            | tcccacccgc |            |            |            |            |      |
| ccaaaccaaa | cctagcctcc | aagagcggga | agaaactaaa | gcaagatagg | ctattaagtg | 5040 |
|            | gaaaatgcct |            |            |            |            |      |
|            | tggccttaat |            |            |            |            |      |
|            | cgagcggtat |            |            |            |            |      |
|            | gcaggaaaga |            |            |            |            |      |
|            | ttgctggcgt |            |            |            |            |      |
|            | agtcagaggt |            |            |            |            |      |
| ccctggaage | tccctcgtgc | geteteetgt | tecgaceetg | ccgcttaccg | gatacctgtc | 5460 |
|            | ccttcgggaa |            |            |            |            |      |
|            | gtcgttcgct |            |            |            |            |      |
|            | ttatccggta |            |            |            |            |      |
|            | gcagccactg |            |            |            |            |      |
| agagttcttg | aagtggtggc | ctaactacgg | ctacactaga | agaacagtat | ttggtatctg | 5760 |
| cgctctgctg | aagccagtta | ccttcggaaa | aagagttggt | agctcttgat | ccggcaaaca | 5820 |
|            | ggtagcggtg |            |            |            |            |      |
|            | gaagateett |            |            |            |            |      |
|            | gggattttgg |            |            |            |            |      |
|            | tgaagtttta |            |            |            |            |      |
|            | ttaatcagtg |            |            |            |            |      |
|            | crcggggggg |            |            |            |            |      |
| ctcataccag | gcctgaatcg | cccatcatc  | cagccagaaa | gtgagggagc | cacggttgat | 6240 |
| gagagetttg | ttgtaggtgg | accagttggt | gattttgaac | ttttgctttg | ccacggaacg | 6300 |
| gtctgcgttg | tcgggaagat | gcgtgatctg | atccttcaac | tcagcaaaag | ttcgatttat | 6360 |

```
tcaacaaagc cgccgtcccg tcaagtcagc gtaatgctct gccagtgtta caaccaatta 6420
accaattctg attagaaaaa ctcatcgagc atcaaatgaa actgcaattt attcatatca 6480
ggattatcaa taccatattt ttgaaaaagc cgtttctgta atgaaggaga aaactcaccq 6540
aggeagttee ataggatgge aagateetgg tateggtetg egatteegae tegteeaaea 6600
tcaatacaac ctattaattt cccctcgtca aaaataaggt tatcaagtga gaaatcacca 6660
tgagtgacga ctgaatccgg tgagaatggc aaaagcttat qcatttcttt ccaqacttqt 6720
tcaacaggcc agccattacg ctcgtcatca aaatcactcg catcaaccaa accqttattc 6780
attegtgatt gegeetgage gagaegaaat aegegatege tgttaaaagg acaattacaa 6840
acaggaatog aatgcaaccg gogcaggaac actgccagcg catcaacaat attttcacct 6900
gaatcaggat attottotaa taootggaat gotgttttoo oggggatogo agtggtgagt 6960
aaccatgcat catcaggagt acggataaaa tgcttgatgg tcggaagagg cataaattcc 7020
gtcagccagt ttagtctgac catctcatct gtaacatcat tggcaacgct acctttgcca 7080
tgtttcagaa acaactetgg cgcatcgggc ttcccataca atcgatagat tgtcqcacct 7140
gattgcccga cattatcgcg agcccattta tacccatata aatcagcatc catgttggaa 7200
tttaatcgcg gcctcgagca agacgtttcc cgttgaatat ggctcataac accccttgta 7260
ttactgttta tgtaagcaga cagttttatt gttcatgatg atatattttt atcttgtgca 7320
atgtaacatc agagattttg agacacaacg tggctttccc cccccccca ttattqaaqc 7380
atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa 7440
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta agaaaccatt 7500
attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg tc
```

<210> 88
<211> 6774
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

## <400> 88

tegegegtht egghgatgae gghgaaaace tehgaeacat geageheeeg gagaegghea 60 cagettgtct gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatqg gtggaqtatt tacqgtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcaq 900 agctogttta gtgaaccgtc agatogcctg gagacgccat ccacqctqtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcqqat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatqg tatagettag cetataggtg tgggttattg accattattg accactecce 1200 tattqqtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac totgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacqccgt cccccqtqcc 1380

| cgcagttttt | attaaacata | gcgtgggatc | tccacgcgaa | tctcgggtac | : gtgttccgga | 1440 |
|------------|------------|------------|------------|------------|--------------|------|
| catgggctct | tctccggtag | cggcggagct | tccacatccg | agccctggtc | ccatgcctcc   | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tggaggccag | , acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | ı gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact | : taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc   | 1740 |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg   | 1800 |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | tectttecat | gggtcttttc   | 1860 |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatogoca   | 1920 |
| ccatgcgcgt | ggagggcatc | cagcgcaact | ggaagcagtg | gtggatctgg | ggcatcctgg   | 1980 |
| gcttctggat | ggtgatgatc | tacaacgtgc | gcggcaacct | gtgggtgacc | gtgtactacg   | 2040 |
| gcgtgcccgt | gtggaaggag | gccaagacca | ccctgttctg | cgccagcgac | gccaaggcct   | 2100 |
| acgacgccga | ggtgcacaac | gtgtgggcca | cccacgcctg | cgtgcccacc | gaccccaacc   | 2160 |
| cccaggagat | ggtgctggag | aacgtgaccg | agaacttcaa | catgtgggag | aacgacatgg   | 2220 |
| tggagcagat | gcaccaggac | atcatcagcc | tgtgggacca | gagcctgaag | ccctgcgtga   | 2280 |
| agctgacccc | cctgtgcgtg | accctgcact | gcagcaaccg | caccatcgac | tacaacaacc   | 2340 |
| gcaccgacaa | catgggcggc | gagatcaaga | actgcagctt | caacatgacc | accgaggtgc   | 2400 |
| gcgacaagcg | cgagaaggtg | cacgccctgt | tctaccgcct | ggacatcgtg | cccctgaaga   | 2460 |
| acgagagcag | caacaccagc | ggcgactacc | gcctgatcaa | ctgcaacacc | agcgccatca   | 2520 |
| cccaggcctg | ccccaaggtg | agcttcgacc | ccatccccat | ccactactgc | gcccccgccg   | 2580 |
| gctacgccat | cctgaagtgc | aacaacaaga | ccttcaacgg | caccggcccc | tgcaacaacg   | 2640 |
| tgagcaccat | ccagtgcacc | cacggcacca | agcccgtggt | gagcacccag | ctgctgctga   | 2700 |
| acggcagcct | ggccgaggag | gagatcatca | tccgcagcaa | gaacctgacc | gacaacqtqa   | 2760 |
| agaccatcat | cgtgcacctg | aacgagagcg | tggagatcaa | ctgcacccgc | cccaacaaca   | 2820 |
| acacccgcaa | gagcatccgc | atcggccccg | gccaggcctt | ctacgccacc | ggcgagatca   | 2880 |
| tcggcgacat | ccgccaggcc | cactgcaaca | tcagccgcac | cgcctggaac | aagaccctqc   | 2940 |
| aggaggtggg | caagaagctg | gccgagcact | tccccaacaa | ggccatcaag | ttcgccaage   | 3000 |
| acagcggcgg | cgacctggag | atcaccaccc | acagcttcaa | ctgccgcggc | gagttcttct   | 3060 |
| actgcaacac | cagcagcctg | ttcaacagca | cctacacccc | caacagcacc | gagaacatca   | 3120 |
| ccggcaccga | gaacagcatc | atcaccatcc | cctgccgcat | caagcagatc | atcaacatgt   | 3180 |
| ggcagggcgt | gggccgcgcc | atgtacgccc | cccccatcga | gggcatcctg | acctgccgca   | 3240 |
| gcaacatcac | cggcctgctg | ctgacccgcg | acggcggcac | cggcatgcac | gacaccgaga   | 3300 |
| tetteegeet | gggcggcggc | gacatgcgcg | acaactggcg | cagcgagctg | tacaagtaca   | 3360 |
| aggtggtgaa | gatcgagccc | ctgggcgtgg | ccccaccaa  | ggccaagctt | accetecage   | 3420 |
| cccgccagct | gctgagcggc | atcgtgcagc | agcagaacaa | cctgctgcgc | gccatcgagg   | 3480 |
| cccagcagca | cctgctgcag | ctgaccgtgt | ggggcatcaa | gcagctgcag | gcccgcatcc   | 3540 |
| tggccgtgga | gcgctacctg | aaggaccagc | agctggagca | gatctggaac | cacaccacct   | 3600 |
| ggatggagtg | ggaccgcgag | atcaacaact | acaccagcct | gatccacagc | ctgatcgagg   | 3660 |
| agagccagaa | ccagcaggag | aagaacgagc | aggagctgct | ggagctggac | aagtgggcca   | 3720 |
| gcctgtggaa | ctggttcaac | atcaccaact | ggctgtggtg | aggatccaga | tctgctgtgc   | 3780 |
| cttctagttg | ccagccatct | gttgtttgcc | cctcccccgt | gccttccttg | accctggaag   | 3840 |
| gtgccactcc | cactgtcctt | tcctaataaa | atgaggaaat | tgcatcgcat | tgtctgagta   | 3900 |
| ggtgtcattc | tattctgggg | ggtggggtgg | ggcaggacag | caagggggag | gattgggaag   | 3960 |
| acaatagcag | gcatgctggg | gatgcggtgg | gctctatggg | tacccaggtg | ctgaagaatt   | 4020 |
| gacccggttc | ctcctgggcc | agaaagaagc | aggcacatcc | ccttctctgt | gacacaccct   | 4080 |
| gtccacgccc | ctggttctta | gttccagccc | cactcatagg | acactcatag | ctcaggaggg   | 4140 |
| ctccgccttc | aatcccaccc | gctaaagtac | ttggagcggt | ctctccctcc | ctcatcagcc   | 4200 |
| caccaaacca | aacctagcct | ccaagagtgg | gaagaaatta | aagcaagata | ggctattaag   | 4260 |
| tgcagaggga | gagaaaatgc | ctccaacatg | tgaggaagta | atgagagaaa | tcatagaatt   | 4320 |
| ttaaggccat | catggcctta | atcttccgct | tcctcgctca | ctgactcgct | gcgctcggtc   | 4380 |
| gttcggctgc | ggcgagcggt | atcagctcac | tcaaaggcgg | taatacggtt | atccacagaa   | 4440 |
| tcaggggata | acgcaggaaa | gaacatgtga | gcaaaaggcc | agcaaaaggc | caggaaccgt   | 4500 |
| aaaaaggccg | cgttgctggc | gtttttccat | aggctccgcc | cccctgacga | gcatcacaaa   | 4560 |
| aatcgacgct | caagtcagag | gtggcgaaac | ccgacaggac | tataaagata | ccaggcgttt   | 4620 |
| ccccctggaa | gctccctcgt | gegeteteet | gttccgaccc | tgccgcttac | cggatacctg   | 4680 |
|            |            |            |            |            |              |      |

```
teegeettte teeetteggg aagegtggeg ettteteata geteaegetg taggtatete 4740
agtteggtgt aggtegtteg etceaagetg ggetgtgtge aegaaceeee egtteageee 4800
gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta 4860
tegecaetgg cageagecae tggtaacagg attageagag egaggtatgt aggeggtget 4920
acagagttet tgaagtggtg geetaactae ggetaeacta gaagaacagt atttggtate 4980
tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa 5040
caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa 5100
aaaggatete aagaagatee titgatetii tetaeggggi etgaegetea giggaaegaa 5160
aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt 5220
ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac 5280
agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc 5340
atagttgcct gactcggggg gggggggcgc tgaggtctgc ctcgtgaaga aggtgttgct 5400
gactcatacc aggcctgaat cgccccatca tocagccaga aagtgaggga gccacggttg 5460
atgagagett tgttgtaggt ggaccagttg gtgattttga acttttgett tgccacggaa 5520
cggtctgcgt tgtcgggaag atgcgtgatc tgatccttca actcagcaaa agttcgattt 5580
attcaacaaa gccgccgtcc cgtcaagtca gcgtaatgct ctgccagtgt tacaaccaat 5640
taaccaattc tgattagaaa aactcatcga gcatcaaatg aaactgcaat ttattcatat 5700
caggattatc aataccatat ttttgaaaaa gccgtttctg taatgaagga gaaaactcac 5760
cgaggcagtt ccataggatg gcaagatect ggtateggte tgcgattecg actegtecaa 5820
catcaataca acctattaat ttcccctcgt caaaaataag gttatcaagt gagaaatcac 5880
catgagtgac gactgaatcc ggtgagaatg gcaaaagctt atgcatttct ttccagactt 5940
gttcaacagg ccagccatta cgctcgtcat caaaatcact cgcatcaacc aaaccgttat 6000
tcattcgtga ttgcgcctga gcgagacgaa atacgcgatc gctgttaaaa ggacaattac 6060
aaacaggaat cgaatgcaac cggcgcagga acactgccag cgcatcaaca atattttcac 6120
ctgaatcagg atattettet aatacetgga atgetgtttt eeeggggate qeaqtqqtqa 6180
gtaaccatgc atcatcagga gtacggataa aatgcttgat ggtcggaaga ggcataaatt 6240
ccgtcagcca gtttagtctg accatctcat ctgtaacatc attggcaacg ctacctttgc 6300
catgtttcag aaacaactct ggcgcatcgg gcttcccata caatcgatag attgtcgcac 6360
etgattgccc gacattatcg cgagcccatt tatacccata taaatcagca tccatgttgg 6420
aatttaatcg cggcctcgag caagacgttt cccgttgaat atggctcata acaccccttg 6480
tattactgtt tatgtaagca gacagtttta ttgttcatga tgatatattt ttatcttgtg 6540
caatgtaaca tcagagattt tgagacacaa cgtggctttc ccccccccc cattattgaa 6600
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 6660
aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca 6720
ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtc
```

```
<210> 89
<211> 6846
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 89
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacattata ttggctcatg 300
tccaacatta ccgccatgtt gacattgat attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta ccggtaaatgg 420
cccgcctggc tgaccgcca acgaccccg cccattgacg tcaataatga cgtatgttcc 480
```

|            |             | ctttccattg |            |             |             |      |
|------------|-------------|------------|------------|-------------|-------------|------|
|            |             | aagtgtatca |            |             |             |      |
|            |             | ggcattatgc |            |             |             |      |
| ttggcagtac | atctacgtat  | tagtcatcgc | tattaccatg | gtgatgcggt  | tttggcagta  | 720  |
| catcaatggg | cgtggatagc  | ggtttgactc | acggggattt | ccaagtctcc  | accccattga  | 780  |
| cgtcaatggg | agtttgttt   | ggcaccaaaa | tcaacgggac | tttccaaaat  | gtcgtaacaa  | 840  |
| ctccgcccca | ttgacgcaaa  | tgggcggtag | gcgtgtacgg | tgggaggtct  | atataagcag  | 900  |
| agctcgttta | gtgaaccgtc  | agatcgcctg | gagacgccat | ccacgctgtt  | ttgacctcca  | 960  |
|            |             | ccagcctccg |            |             |             |      |
|            |             | taagtaccgc |            |             |             |      |
|            |             | ttttggcttg |            |             |             |      |
|            |             | cctataggtg |            |             |             |      |
|            |             | attactaatc |            |             |             |      |
|            |             | tetgteette |            |             |             |      |
|            |             | tttacaaatt |            |             |             |      |
|            |             | gcgtgggatc |            |             |             |      |
|            |             | caacaaaacr |            |             |             |      |
|            |             | cagctccttg |            |             |             |      |
|            |             | cagtgtgccg |            |             |             |      |
| gaaaatgaga | atagagatta  | ggctcgcacg | actalogoog | atoragonast | taaaacaaca  | 1690 |
| gaaaacgagc | gragagarra  | ctgagttgtt | geegaegeag | angatona    | catagetageg | 1740 |
| gcagaagaag | tattanggrag | ggagggcagt | gtactctgac | cagtactcat  | tactaccaca  | 1900 |
| grigeggige | cgccaacggc  | ctgacagact | grageregag | techtteest  | gatattt     | 1960 |
| besset     | gacataatag  | cegacagace | antabaccac | coccecae    | gggtttttt   | 1000 |
|            |             | cgtgtgatca |            |             |             |      |
| ccatgegegt | ggagggeate  | cagcgcaact | ggaagcagcg | gtggatttgg  | ggeaceetgg  | 1300 |
|            |             | tacaacgtgc |            |             |             |      |
|            |             | gccaagacca |            |             |             |      |
|            |             | gtgtgggcca |            |             |             |      |
|            |             | aacgtgaccg |            |             |             |      |
|            |             | atcatcagcc |            |             |             |      |
| agetgaeece | cetgtgegtg  | accetgeact | geageaaceg | caccatcgac  | tacaacaacc  | 2340 |
|            |             | gagatcaaga |            |             |             |      |
|            |             | cacgccctgt |            |             |             |      |
| acgagagcag | caacaccagc  | ggcgactacc | geetgateaa | etgeaacacc  | agegeeatea  | 2520 |
|            |             | agcttcgacc |            |             |             |      |
|            |             | aacaacaaga |            |             |             |      |
|            |             | cacggcacca |            |             |             |      |
|            |             | gagatcatca |            |             |             |      |
|            |             | aacgagagcg |            |             |             |      |
|            |             | atcggccccg |            |             |             |      |
|            |             | cactgcaaca |            |             |             |      |
|            |             | gccgagcact |            |             |             |      |
|            |             | atcaccaccc |            |             |             |      |
|            |             | ttcaacagca |            |             |             |      |
|            |             | atcaccatcc |            |             |             |      |
|            |             | atgtacgccc |            |             |             |      |
|            |             | ctgacccgcg |            |             |             |      |
|            |             | gacatgcgcg |            |             |             |      |
|            |             | ctgggcgtgg |            |             |             |      |
|            |             | atcgtgcagc |            |             |             |      |
|            |             | ctgaccgtgt |            |             |             |      |
|            |             | aaggaccagc |            |             |             |      |
|            |             | atcaacaact |            |             |             |      |
|            |             | aagaacgagc |            |             |             |      |
| gcctgtggaa | ctggttcaac  | atcaccaact | ggctgtggta | catcaagctg  | ttcatcatga  | 3780 |

|            |            |            |            |            | tgaggatcca |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | gtgccttcct |      |
| tgaccctgga | aggtgccact | cccactgtcc | tttcctaata | aaatgaggaa | attgcatcgc | 3960 |
| attgtctgag | taggtgtcat | tctattctgg | ggggtggggt | ggggcaggac | agcaaggggg | 4020 |
| aggattggga | agacaatagc | aggcatgctg | gggatgcggt | gggctctatg | ggtacccagg | 4080 |
| tgctgaagaa | ttgacccggt | tcctcctggg | ccagaaagaa | gcaggcacat | ccccttctct | 4140 |
| gtgacacacc | ctgtccacgc | ccctggttct | tagttccagc | cccactcata | ggacactcat | 4200 |
|            |            |            |            |            | gtctctccct |      |
|            |            |            |            |            | taaagcaaga |      |
|            |            |            |            |            | taatgagaga |      |
|            |            |            |            |            | cactgactcg |      |
|            |            |            |            |            | ggtaatacgg |      |
|            |            |            |            |            | ccagcaaaag |      |
|            |            |            |            |            | ccccctgac  |      |
|            |            |            |            |            | actataaaga |      |
|            |            |            |            |            | cctgccgctt |      |
|            |            |            |            |            | tagctcacgc |      |
|            |            |            |            |            | gcacgaaccc |      |
|            |            |            |            |            | caacccggta |      |
|            |            |            |            |            | agcgaggtat |      |
|            |            |            |            |            | tagaagaaca |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | tggtagctct |      |
|            |            |            |            |            | gcagcagatt |      |
|            |            |            |            |            | gtctgacgct |      |
|            |            |            |            |            | aaggatette |      |
|            |            |            |            |            | atatgagtaa |      |
|            |            |            |            |            | gatctgtcta |      |
|            |            |            |            |            | gcctcgtgaa |      |
|            |            |            |            |            | gaaagtgagg |      |
|            |            |            |            |            | gaacttttgc |      |
|            |            |            |            |            | caactcagca |      |
|            |            |            |            |            | ctctgccagt |      |
|            |            |            |            |            | tgaaactgca |      |
|            |            |            |            |            | tgtaatgaag |      |
|            |            |            |            |            | tetgegatte |      |
|            |            |            |            |            | aggttatcaa |      |
|            |            |            |            |            | ttatgcattt |      |
|            |            |            |            |            | ctcgcatcaa |      |
|            |            |            |            |            | tcgctgttaa |      |
|            |            |            |            |            | agcgcatcaa |      |
|            |            |            |            |            | ttcccgggga |      |
|            |            |            |            |            | atggtcggaa |      |
|            |            |            |            |            | tcattggcaa |      |
|            |            |            |            |            | tacaatcgat |      |
|            |            |            |            |            | tataaatcag |      |
|            |            |            |            |            | atatggctca |      |
| taacacccct | tgtattactg | tttatgtaag | cagacagttt | tattgttcat | gatgatatat | 6600 |
| ttttatcttg | tgcaatgtaa | catcagagat | tttgagacac | aacgtggctt | tecceccee  | 6660 |
| cccattattg | aagcatttat | cagggttatt | gtctcatgag | cggatacata | tttgaatgta | 6720 |
| tttagaaaaa | taaacaaata | ggggttccgc | gcacatttcc | ccgaaaagtg | ccacctgacg | 6780 |
|            |            |            |            |            | acgaggccct |      |
| ttcgtc     |            |            |            |            |            | 6846 |
|            |            |            |            |            |            |      |

<210> 90
<211> 6853
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccqc atcagattqq 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgit gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcqtt acataactta cqqtaaatqq 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acqtcaatqq qtqqaqtatt tacqqtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataaqcaq 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetqtt ttqaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tettatgeat getatactgt tittggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeaq aettaggeac 1560 agcacaatgo ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980 gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040 gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100 acgacgecga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160 cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtqqqaq aacqacatqq 2220 tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agetgacece cetgtgegtg accetgeact geageaaceg caccategae tacaacaace 2340 gcaccgacaa catgggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400 gcgacaagcg cgagaaggtg cacgccctgt tctaccgcct ggacatcgtg cccctgaaga 2460 acgagageag caacaccage ggcgactace geetgateaa etgcaacace agegeeatea 2520 cccaggcctg ccccaaggtg agcttcgacc ccatccccat ccactactgc gcccccgccg 2580 gctacgccat cctgaagtgc aacaacaaga ccttcaacgg caccggcccc tgcaacaacg 2640 tgagcaccat ccagtgcacc cacggcacca agcccgtggt gagcacccag ctgctgctga 2700 acggcagcct ggccgaggag gagatcatca tccgcagcaa gaacctgacc gacaacgtga 2760

| agaccatcat  | cgtgcacctg | aacgagagcg | tggagatcaa | ctgcacccgc  | cccaacaaca  | 2820 |
|-------------|------------|------------|------------|-------------|-------------|------|
|             | gagcatccgc |            |            |             |             |      |
| toggogacat  | ccgccaggcc | cactgcaaca | tcagccgcac | cgcctggaac  | aagaccctgc  | 2940 |
|             | caagaagctg |            |            |             |             |      |
|             | cgacctggag |            |            |             |             |      |
|             | cagcagcctg |            |            |             |             |      |
|             | gaacagcatc |            |            |             |             |      |
|             | gggccgcgcc |            |            |             |             |      |
|             | cggcctgctg |            |            |             |             |      |
|             | cgagggcggc |            |            |             |             |      |
|             | gatcaagccc |            |            |             |             |      |
|             | gctgagcggc |            |            |             |             |      |
| cacaccage   | catgctgcag | ctaeccatat | ageagageaa | acaactacaa  | according   | 3540 |
| tagagetagea | gegetacetg | ccgaccgcgc | agggcaecaa | catctooaac  | aacatcacct  | 3600 |
| cggccarcga  | ggaccgcgag | atasaasaat | agecegagga | catctaggaac | ctactageaga | 3660 |
|             |            |            |            |             |             |      |
| acagccagaa  | ccagcaggag | aagaacgagc | aggacccgcc | ggccccggac  | ttastassas  | 3700 |
|             | ctggttcggc |            |            |             |             |      |
| rcaraaacaa  | cctgatcggc | etgegeatea | cecegeege  | getgageate  | cgaacacgtg  | 3840 |
|             | ctgctgtgcc |            |            |             |             |      |
|             | ccctggaagg |            |            |             |             |      |
|             | gtctgagtag |            |            |             |             |      |
|             | attgggaaga |            |            |             |             |      |
|             | tgaagaattg |            |            |             |             |      |
|             | acacaccctg |            |            |             |             |      |
|             | tcaggagggc |            |            |             |             |      |
|             | tcatcagccc |            |            |             |             |      |
|             | gctattaagt |            |            |             |             |      |
|             | catagaattt |            |            |             |             |      |
|             | cgctcggtcg |            |            |             |             |      |
|             | tccacagaat |            |            |             |             |      |
|             | aggaaccgta |            |            |             |             |      |
|             | catcacaaaa |            |            |             |             |      |
|             | caggcgtttc |            |            |             |             |      |
| gccgcttacc  | ggatacctgt | ccgcctttct | cccttcggga | agcgtggcgc  | tttctcatag  | 4800 |
|             | aggtatctca |            |            |             |             |      |
| cgaacccccc  | gttcagcccg | accgctgcgc | cttatccggt | aactatcgtc  | ttgagtccaa  | 4920 |
| cccggtaaga  | cacgacttat | cgccactggc | agcagccact | ggtaacagga  | ttagcagagc  | 4980 |
| gaggtatgta  | ggcggtgcta | cagagttctt | gaagtggtgg | cctaactacg  | gctacactag  | 5040 |
| aagaacagta  | tttggtatct | gcgctctgct | gaagccagtt | accttcggaa  | aaagagttgg  | 5100 |
| tagctcttga  | tccggcaaac | aaaccaccgc | tggtagcggt | ggtttttttg  | tttgcaagca  | 5160 |
| gcagattacg  | cgcagaaaaa | aaggatctca | agaagatcct | ttgatctttt  | ctacggggtc  | 5220 |
|             | tggaacgaaa |            |            |             |             |      |
|             | tagateettt |            |            |             |             |      |
|             | tggtctgaca |            |            |             |             |      |
|             | cgttcatcca |            |            |             |             |      |
|             | ggtgttgctg |            |            |             |             |      |
|             | ccacggttga |            |            |             |             |      |
|             | gccacggaac |            |            |             |             |      |
|             | gttcgattta |            |            |             |             |      |
|             | acaaccaatt |            |            |             |             |      |
|             | tattcatatc |            |            |             |             |      |
|             | aaaactcacc |            |            |             |             |      |
|             | ctcgtccaac |            |            |             |             |      |
|             | agaaatcacc |            |            |             |             |      |
|             | tccagacttg |            |            |             |             |      |
|             |            |            |            |             | • •         | -    |

```
gcatcaacca aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg 6120 ctgttaaaag gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc 6180 gcatcaacaa tattttcacc tgaatcagga tattcttcta atacctggaa tgctgttttc 6240 ccgggggatcg cagtggtgag taaccatgca tcatcaggag tacggataaa atgcttgatg 6300 gtcggaagag gcataaattc cgtcagccag tttagtctga ccatctcatc tgtaacatca 6360 ttggcaacgc tacctttgcc atgtttcaga aacaactctg gcgcatcggg cttcccatac 6420 aatcgataga ttgtcgcacc tgattgcccg acattatcgc gagcccattt atacccatat 6480 agatatattt tatcttgtga aatgtaacat cagagattt gggcaccagg ccgttgaata 6540 ttggctcataa cacccttgt attactgtt atgtaagcag acagttttat tgttcatgat 6600 gatatatttt tatcttgtgc aatgtaacat cagagatttt gagacacaac gtggctttcc 6660 ccccccccc attattgag catttatcag ggttattgtc tcatgagcgg atacatattt 6720 gaatgtattt agaaaaataa acaaataggg gttccgcga catttccccg aaaagtgcca 6840 aggccctttc gtc
```

<210> 91 <211> 6761 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes

<400> 91 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttqqcqqqtq tcqqqqctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta eegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tqcccacttq qcaqtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tottatgcat gotatactgt ttttggottg gggodtatad accodegett cottatgcta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaeteeec 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cqcaqttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560 aqcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcaqaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

| attacaatac  | tattaacaat | agagggcagt  | gtagtctgag | cantactont | tactaccasa | 1000 |
|-------------|------------|-------------|------------|------------|------------|------|
| cacaccacca  | gacataatag | ctgacagact  | aacagactgt | teettieest | gggtgttgt  | 1060 |
| tacaatcacc  | atcatcaaca | catatastas  | gatatcgcgg | ccactctage | gagettette | 1000 |
| ccatgcgcgt  | acacaacatc | cadaccadet  | ggcagaacct | ataccactca | gacaccycca | 1000 |
| tectagaeat  | actagtasta | tacagegee   | ccgagaacct | gtggtgttgg | gycaccatga | 1300 |
| acatacccat  | graggegaee | accasaccs   | ccctgttctg | gegggeggee | graceacg   | 2040 |
| acdacaccda  | actacacaac | geegagaeea  | cccacgcctg | catacacaca | gccaaggccc | 2100 |
| cccagaaga   | ggegeacaac | gegegggaga  | accacgeetg | astatasaaa | gaccccaacc | 2160 |
| tagaagaa    | ggacccggag | atcatcacc   | aggacttcaa | catgtggege | aacaacacgg | 2220 |
| agetgaecee  | cctatacata | accatcaged  | tgtgggacca | gageetgaag | ccctgcgtga | 2280 |
| agatgaccca  | ctacaactta | accetggact  | gcaacgccac | cgccagcaac | grgaecaacg | 2340 |
| acacctatt   | chacaagete | gacatcatca  | ccgagctgaa | ggacaagaag | cagcaggigi | 2400 |
| accocctoat  | caactageeg | accaccacca  | agatcaacga | gaagaacyag | accgacaage | 2460 |
| accepted    | catccactac | tacagegeea  | tcacccaggc | cogcoccaag | grgagerreg | 2520 |
| ccaacttcaa  | caccaccac  | cactagaaga  | ccggcttcgc | catectgaag | rgcaaggaca | 2580 |
| tecacceat   | gatcagge   | cacctactac  | acgtgagcac | egtgeagtge | acccacggca | 2640 |
| agatocacaa  | cascascata | aggregates  | tgaacggcag | cctggccgag | gagggcatcc | 2700 |
| contragant  | caactacaca | accaacaacg  | ccaagaccat | categugeag | ctggacaagg | 2/60 |
| ccgcgaagac  | caactycact | egeeccaaca  | acaacacccg | caagggcgtg | cgcatcggcc | 2820 |
| Regterage   | ccccacgee  | acceggeggea | tcatcggcga | catccgccag | gcccactgcc | 2880 |
| acgigageeg  | cgccaagtgg | aacgacacce  | tgcgcggcgt | ggccaagaag | ctgcgcgagc | 2940 |
| cccacacatt  | caagaccatc | accettegaga | agagcagcgg | cggcgacatc | gagatcacca | 3000 |
| genetagett  | caccegegge | ggcgagttet  | tctactgcaa | caccagegge | ctgttcaaca | 3060 |
| tancetagga  | gagcaacagc | accgagagea  | acaacaccac | cagcaacgac | accatcaccc | 3120 |
| coacceact   | caccaagcag | attattatta  | tgtggcagaa | ggrgggccag | gccatgtacc | 3180 |
| cccccccat   | ccagggegeg | accegetgeg  | agagcaacat | caccggcctg | ctgctgaccc | 3240 |
| gegaeggegg  | caacaacage | accaacgaga  | tetteegeee | caacaacaac | aacatgcgcg | 3300 |
| acaactggcg  | cagegagetg | tacaagtaca  | aggtggtgaa | gatcgagccc | ctgggcgtgg | 3360 |
| ceeeeageeg  | cgccaagett | accgcccagg  | cccgccagct | gctgagcggc | atcgtgcagc | 3420 |
| agcagagcaa  | cetgetgege | gecategagg  | cccagcagca | catgorgaag | ctgaccgtgt | 3480 |
| aggigeaceaa | geagetgeag | geeegegege  | tggccgtgga | gcgctacctg | aaggaccagc | 3540 |
| agetegagat  | ctgggacaac | atgacctgge  | tgcagtggga | caaggagatc | agcaactaca | 3600 |
| acctactac   | ccacaacctg | accgaggaga  | gccagaacca | gcaggagaag | aacgagcagg | 3660 |
| tataataaaa  | cccggacaag | rgggecagee  | tgtggaactg | gttcgacatc | agccgctggc | 3720 |
| tacasatasa  | acgraggate | cagatetget  | gtgccttcta | gttgccagcc | atctgttgtt | 3780 |
| tanastanaa  | asabbasaba | cutgaceetg  | gaaggtgcca | ccccactgt  | cctttcctaa | 3840 |
| caaaacgagg  | adactgcate | geattgeetg  | agtaggtgtc | attetattet | aaaaaaraaa | 3900 |
| gragggragg  | teageaaygg | ggaggartgg  | gaagacaata | gcaggcatgc | tggggatgcg | 3960 |
| gragagerera | cyggcaccca | ggtgctgaag  | aattgacccg | greereerg  | ggccagaaag | 4020 |
| aagcaggcac  | baggaggete | ctgtgacaca  | ccctgtccac | gcccctggtt | cttagttcca | 4080 |
| gtacttacca  | caggacacce | atageteagg  | agggctccgc | cttcaatccc | accegetaaa | 4140 |
| gtacttggag  | attacages  | achagatat   | agcccaccaa | accaaaccta | gcctccaaga | 4200 |
| gragaagaa   | actaaagcaa | gataggetat  | taagtgcaga | gggagagaaa | atgeetecaa | 4260 |
| cacgegagga  | tteractter | tagetaleta  | aattttaagg | ccatgattta | aggccatcat | 4320 |
| ggccccaatc  | agatasatas | cogocoaccg  | actcgctgcg | cccggccgcc | cggctgcggc | 4380 |
| gageggeate  | agercaerca | aaggeggeaa  | tacggttatc | cacagaacca | ggggataacg | 4440 |
| taggaaagaa  | thtashas   | aaaggeeage  | aaaaggccag | gaaccgtaaa | aaggccgcgt | 4500 |
| charageata  | ccccacagg  | cceegeeeee  | ctgacgagca | tcacaaaaat | cgacgctcaa | 4560 |
| gccagaggcg  | gegaaaeeeg | acaggactat  | aaagatacca | ggcgtttccc | cctggaagct | 4620 |
| eccuegueg   | ctctcctgtt | ccgaccccgc  | cgcttaccgg | atacctgtcc | gcctttctcc | 4680 |
| backtoost=  | cgtggcgctt | teteataget  | cacgctgtag | gratctcagt | tcggtgtagg | 4740 |
| batacastas  | caagetggge | cgtgtgcacg  | aacccccgt  | ccagcccgac | cgctgcgcct | 4800 |
| cacceggeaa  | ccatcgtctt | gagtccaacc  | cggtaagaca | cgacttatcg | ccactggcag | 4860 |
| caycoaccgg  | caacaggatt | agcagagcga  | ggtatgtagg | cggtgctaca | gagttcttga | 4920 |
| agragacht   | caactacggc | tacactagaa  | gaacagtatt | tggtatctgc | gctctgctga | 4980 |
| agccagetac  | ecceggaaaa | agagttggta  | gctcttgatc | cggcaaacaa | accaccgctg | 5040 |

```
gtageggtgg tttttttgtt tgcaageage agattaegeg cagaaaaaaa ggatetcaag 5100
aagateettt gatetttet aeggggtetg aegeteagtg gaacgaaaac teaegttaag 5160
ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat 5220
gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct 5280
taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata gttgcctgac 5340
teggggggg ggggcgctga ggtetgeete gtgaagaagg tgttgetgae teataceagg 5400
cetgaatege cecateatee agecagaaag tgagggagee aeggttgatg agagetttgt 5460
cgggaagatg cgtgatctga tccttcaact cagcaaaagt tcgatttatt caacaaagcc 5580
geogtecegt caagtcageg taatgetetg coagtgttac aaccaattaa ccaattetga 5640
ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat 5700
accatatttt tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca 5760
taggatggca agatectggt ateggtetge gatteegaet egteeaacat caatacaace 5820
tattaatttc ccctcgtcaa aaataaggtt atcaagtgag aaatcaccat gagtgacgac 5880
tgaatccggt gagaatggca aaagcttatg catttctttc cagacttgtt caacaggcca 5940
gccattacgc tcgtcatcaa aatcactcgc atcaaccaaa ccgttattca ttcgtgattg 6000
cgcctgagcg agacgaaata cgcgatcgct gttaaaagga caattacaaa caggaatcga 6060
atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaqqata 6120
ttcttctaat acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcatc 6180
atcaggagta cggataaaat gcttgatggt cggaagaggc ataaattccg tcagccagtt 6240
tagtctgacc atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa 6300
caactctggc gcatcgggct tcccatacaa tcgatagatt gtcgcacctg attgcccgac 6360
attatcgcga gcccatttat acccatataa atcagcatcc atgttggaat ttaatcgcgg 6420
cetegageaa gacgtttccc gttgaatatg gctcataaca cecettgtat tactgtttat 6480
gtaagcagac agttttattg ttcatgatga tatattttta tcttgtgcaa tgtaacatca 6540
gagattttga gacacaacgt ggctttcccc cccccccat tattgaagca tttatcaggg 6600
ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt 6660
tecgegeaca tttececegaa aagtgeeace tgaegtetaa gaaaceatta ttateatgae 6720
attaacctat aaaaataggc gtatcacgag gccctttcgt c
                                                                6761
```

<210> 92 <211> 6836 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 92 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 tiggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tittggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840

| ctccgcccca  | ttgacgcaaa | tgggcggtag                               | gcgtgtacgg | taggaggtct  | atataaqcaq   | 900  |
|-------------|------------|--|------------|-------------|--------------|------|
| agctcgttta  | gtgaaccgtc | agategeetg                               | gagacgccat | ccacactatt  | ttgacctcca   | 960  |
| tagaagacac  | cgggaccgat | ccagcctccq                               | cggccgggaa | cagtacatta  | gaacgcggat   | 1020 |
| teceegtgee  | aagagtgacg | taagtaccgc                               | ctatagactc | tataggcaca  | ccctttaac    | 1080 |
| tcttatgcat  | gctatactgt | ttttqqcttq                               | gggcctatac | accccccctt  | ccttatecta   | 1140 |
| taggtgatgg  | tatagettag | cctataggtg                               | tgggttattg | accattatto  | accachece    | 1200 |
| tattggtgac  | gatactttcc | attactaatc                               | cataacatgg | ctctttacca  | caactatctc   | 1260 |
| tattggctat  | atqccaatac | tctqtccttc                               | agagactgac | acqqactctq  | tatttttaca   | 1320 |
| ggatggggtc  | ccatttatta | tttacaaatt                               | cacatataca | acaacaccat  | ccccataca    | 1380 |
| cqcaqttttt  | attaaacata | acataggato                               | tccacgcgaa | tctcagatec  | atattagaga   | 1440 |
| catgggctct  | tctccqqtaq | caacaaaact                               | tccacatccg | accetacte   | ccatocctcc   | 1500 |
| agcggctcat  | gatcactcaa | cagctccttg                               | ctcctaacag | tagacaccaa  | acttagggag   | 1560 |
| agcacaatgc  | ccaccaccac | cagtatacca                               | cacaaggccg | tagaagacaa  | atetatatatat | 1620 |
| gaaaatgagc  | gtggagattg | aactcacaca                               | gctgacgcag | atggaagact  | taaggggcc    | 1600 |
| gcagaagaag  | atgcaggcag | ctgagttgtt                               | gtattctgat | aagagtcaga  | caaggcagcg   | 1740 |
| attacaatac  | tottaacoot | agagggagt                                | gtagtctgag | cagtactcat  | tactaccaca   | 1000 |
| cgcgccacca  | gacataatag | chacagact                                | aacagactgt | techtrocat  | cactatttta   | 1000 |
| tgcagtcacc  | gtcgtcgaca | catatastas                               | gatatcgcgg | cccctcccat  | gggccccccc   | 1000 |
| ccatacacat  | gcacaacatc | cadaccadct                               | ggcagaacct | atacacta    | gacacegeca   | 1920 |
| tectaggest  | actagtasta | tacaccacca                               | ccgagaacct | gragegergg  | ggcaccatga   | 1980 |
| acatacccat  | atagaagac  | accagagaga                               | ccctgttctg | graggraggee | grgractacg   | 2040 |
| acgacaccga  | goggaaggac | gtatagacca                               | ccctgccctg | cgccagegae  | gccaaggcct   | 2100 |
| cccagaggt   | ggegededae | accetoacce                               | cccacgcctg | cgtgcccace  | gaccccaacc   | 2160 |
| tagaggagat  | ggacccggag | atcatcage                                | aggacttcaa | catgeggege  | aacaacatgg   | 2220 |
| agetgacece  | cctatacata | accatcagee                               | tgtgggacca | gageetgaag  | ccctgcgtga   | 2280 |
| agetgacee   | ctagagetta | accordgact                               | gcaacgccac | cgccagcaac  | grgaccaacg   | 2340 |
| agacgcgcaa  | ctgcagette | aacaccacca                               | ccgagctgaa | ggacaagaag  | cagcaggtgt   | 2400 |
| acageetget  | ccacaagetg | gacgeggege                               | agatcaacga | gaagaacgag  | accgacaagt   | 2460 |
| accected    | catcagetae | bassassassassassassassassassassassassass | tcacccaggc | ctgccccaag  | gtgagetteg   | 2520 |
| ccaacttcaa  | cacccaccac | cgcgcccccg                               | ccggcttcgc | cateetgaag  | tgcaaggaca   | 2580 |
| tecareceat  | capcaccage | ccctgcaaga                               | acgtgagcac | cgtgeagtge  | acccacggca   | 2640 |
| adatccccac  | gaccagcacc | cagetgetge                               | tgaacggcag | ceuggeegag  | gagggcatcc   | 2700 |
| ccatasaast  | caactacacc | accaacaacg                               | ccaagaccat | categegeag  | ctggacaagg   | 2760 |
| ccgcgaagac  | cttctaccc  | acconggggg                               | acaacacccg | caagggcgtg  | cgcatcggcc   | 2820 |
| acataaacca  | caccaaataa | accegegea                                | tcatcggcga | catcegecag  | gcccactgcc   | 2880 |
| acttcaacaa  | caccaagigg | alcyacaccc                               | tgcgcggcgt | ggccaagaag  | cracacaaac   | 2940 |
| cccacaagaa  | catgaccacc | accuregaga                               | agagcagcgg | cggcgacate  | gagaccacca   | 3000 |
| acacctagac  | gaggaagag  | ggegageeee                               | tctactgcaa | caccagegge  | ctgttcaaca   | 3060 |
| taacetacca  | gagcaacagc | accgagagea                               | acaacaccac | cagcaacgac  | accatcaccc   | 3120 |
| cccccccc    | caccaagcag | attattaata                               | tgtggcagaa | ggtgggccag  | gccatgtacc   | 3180 |
| acaecacaa   | ccagggegeg | accegeracy                               | agagcaacat | caccggcctg  | ctgctgaccc   | 3240 |
| acaactooco  | caacaacage | tacaacgaga                               | tetteegeee | cggcggcggc  | aacatgcgcg   | 3300 |
| ccccaacca   | cagcgagety | accasguaca                               | aggtggtgaa | gategageee  | ctgggcgtgg   | 3360 |
| accacagoog  | cotcotcoc  | accycccagg                               | cccgccagct | gctgagcggc  | atcgtgcagc   | 3420 |
| agcagagcaa  | ccegcegege | gecategagg                               | cccagcagca | catgetgaag  | ctgaccgtgt   | 3480 |
| agetegaget  | gragergrag | stangetasa                               | tggccgtgga | gcgctacctg  | aaggaccagc   | 3540 |
| agecegagae  | ctgggacaac | atgacctggc                               | tgcagtggga | caaggagatc  | agcaactaca   | 3600 |
| acetactecae | ctacaacetg | accgaggaga                               | gccagaacca | gcaggagaag  | aacgagcagg   | 3660 |
| tataataat   | couggacaag | rgggccagec                               | tgtggaactg | gttcgacatc  | agccgctggc   | 3720 |
| traccatact  | caagaccccc | tacacgaceg                               | tgggcggcct | gatcggcctg  | cgcatcgtgt   | 3780 |
| cagagatata  | gagegegate | cyaacacgtg                               | ggatccagat | ctgctgtgcc  | ttctagttgc   | 3840 |
| agecateeg   | cegeeegeee | eccecegeg                                | ccttccttga | ccctggaagg  | tgccactccc   | 3900 |
| attotocas   | cccaacaaaa | rgaggaaact                               | gcatcgcatt | gtctgagtag  | gtgtcattct   | 3960 |
| catactors   | 9-99995999 | gcaggacagc                               | aaggggagg  | attgggaaga  | caatagcagg   | 4020 |
| tectages    | argeggeggg | coccacgggt                               | acccaggtgc | cgaagaattg  | acceggttee   | 4080 |
| ccccgggcca  | yaaagaagca | ggcacatccc                               | cttctctgtg | acacaccctg  | tccacgcccc   | 4140 |

```
tggttettag ttccagecce acteatagga cacteatage tcaggaggge tccgccttca 4200
atoccacceg ctaaagtact tggageggte tetecetece teateagece accaaaceaa 4260
acctagcctc caagagtggg aagaaattaa agcaagatag gctattaagt gcagagggag 4320
agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat catagaattt taaggccatg 4380
atttaaggcc atcatggcct taatcttccg cttcctcgct cactgactcg ctgcgctcgg 4440
tegttegget geggegageg gtateagete acteaaagge ggtaataegg ttatecacag 4500
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 4560
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccctqac qaqcatcaca 4620
aaaatcgacg Ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcqt 4680
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 4740
tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 4800
tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc 4860
ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact 4920
tategecaet ggeageagee actggtaaca ggattageag agegaggtat gtaggeggtg 4980
ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca qtatttqqta 5040
tetgegetet getgaageea gttacetteg gaaaaagagt tggtagetet tgateeggea 5100
aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa 5160
aaaaaggato tcaagaagat cetttgatet tttetaeggg gtetgaeget cagtggaaeg 5220
aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc 5280
ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg 5340
acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat 5400
ccatagttgc ctgactcggg ggggggggc gctgaggtct gcctcgtgaa gaaggtgttg 5460
ctgactcata ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt 5520
tgatgagagc tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacgg 5580
aacggtetge gttgteggga agatgegtga tetgateett caacteagea aaagttegat 5640
ttattcaaca aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca 5700
attaaccaat tetgattaga aaaacteate gagcateaaa tgaaactgea atttatteat 5760
atcaggatta tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc 5820
accgaggcag ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc 5880
aacatcaata caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc 5940
accatgagtg acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac 6000
ttgttcaaca ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt 6060
attcattcgt gattqcqcct qaqcqaqacq aaatacqcqa tcqctqttaa aaqqacaatt 6120
acaaacagga atcgaatgca accggcgcag gaacactgcc agcgcatcaa caatattttc 6180
acctgaatca ggatattett etaataeetg gaatgetgtt tteeegggga tegeagtggt 6240
gagtaaccat gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa 6300
ttccgtcagc cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt 6360
gccatgtttc agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc 6420
acctgattgc ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt 6480
ggaatttaat cgcggcctcg agcaagacgt ttcccgttga atatggctca taacacccct 6540
tgtattactg tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg 6600
tgcaatgtaa catcagagat tttgagacac aacgtggctt tcccccccc cccattattg 6660
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 6720
taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac 6780
cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc
```

```
<210> 93
<211> 6800
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
```

<400> 93 tegegegttt eggtgatgae ggtgaaaaee tetgacaeat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta ceqecatqtt qacattqatt attgactagt tattaatagt aatcaattae 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtetec aceccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 etcegececa ttgaegeaaa tgggeggtag gegtgtaegg tgggaggtet atataageag 900 ageteghtha ghgaaceghe agategeeng gagaegeeat ceaegenght thigaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt tettggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactitcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagaetgae aeggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeg agecetggte coatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgago gtggagattg ggotogoacg gotgaogoag atggaagaot taaggoagog 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgcgt gaaggagacc cagatgaact ggcctaacct gtggaagtgg ggcatcctga 1980 tectgggeet ggtgateate tgeagegeea gegaeaacet gtgggtgaee gtgtaetaeg 2040 gegtgeeegt gtggegegae geegaeacea ceetgttetg egeeagegae geeaaggeee 2100 acgagaccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160 cccaggaggt ggacctggag aacatcaccg agaacttcaa catgtggaag aacaacatgg 2220 tggagcagat gcaggaggac gtgatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agetgaeece cetgtgegtg accetgaact geatcaaege caacetgaee aaegtgaaea 2340 acatcateca eggececaac atcateggea acatcacega egaggtgege aactgeaget 2400 tcaacatgac caccgagate cgcgacaaga agcagaaggt gcacgccctg ttctacaagc 2460 tggacctggt gcagatcgag gacaagaaca gcagcgagta ccgcctgatc aactgcaaca 2520 ccagcgtgat caagcaggcc tgccccaaga tcagcttcga ccccatcccc atccactact 2580 gcacccccgc cggctacgcc atcctgaagt gcaacgacaa gaacttcaac ggcaccggcc 2640 cctgcaccaa cgtgagcagc gtgcagtgca cccacggcat caaqcccqtq gtgagcaccc 2700 agetgetget gaacggcage etggecgagg aggagateat catecgcage gagaacetga 2760 ccaacaacgc caagaccatc atcgtgcacc tgaacaagag cgtggagatc aactgcaccc 2820 geoccageaa caacaccege accageacca ceateggeee eggecaggtg ttetaccgea 2880 ccggcgacat caccggcaac atccgcaagg cctactgcga gatcaacggc accaagtgga 2940 acgaggeet gaagcaggtg accaagaage tgaaggagea etteaacaac aagaccatea 3000 tettecagee ecceagegge ggegacetgg agateaceat geaceaette aactgeegeg 3060 gcgagttctt ctactgcaac accacccgcc tgttcaacaa cacctgcatc ggcaacgaga 3120 ccateggegg etgeaacgac accateatee tgeeetgeaa gateaageag ateateaaca 3180

| +=+======= |             |            |            |            |            |      |
|------------|-------------|------------|------------|------------|------------|------|
| tgcggcaggg | egeeggeetg  | gccatgtacg | ccccccat   | cagcggccgc | atcaactgcg | 3240 |
| Lyageaacat | caccggcatc  | ctgctgaccc | gcgacggcgg | cgccaacaac | accagcaacg | 3300 |
| agaccccccg | ccccggcggc  | ggcaacatca | aggacaactg | gcgcagcgag | ctgtacaagt | 3360 |
| acaaggtggt | gcagatcgag  | cccctgggca | tegeceecae | ccgcgccaag | cttaccgtgc | 3420 |
| aggcccgcca | gctggtgagc  | ggcatcgtgc | agcagcagag | caacctgctg | cgcgccatcg | 3480 |
| aggeceagea | gcacctgctg  | cagctgaccg | tgtggggcat | caagcagctg | caggcccgcg | 3540 |
| ractaaccat | ggagcgctac  | ctgaaggacc | agaagtteet | cgaggagatc | tggagcaaca | 3600 |
| tgacctggat | cgagtgggag  | cgcgagatca | gcaactacac | caaccagatc | tacgagatcc | 3660 |
| tgaccgagag | ccagaaccag  | caggaccgca | acgagaagga | cctgctggag | ctggacaagt | 3720 |
| gggcgagcct | gtggagctgg  | ttcgacatca | ccaactggct | gtggtgaaca | cgtgggatcc | 3780 |
| agatetgetg | tgccttctag  | ttgccagcca | tctgttgttt | gcccctcccc | cgtgccttcc | 3840 |
| ttgaccctgg | aaggtgccac  | tcccactgtc | ctttcctaat | aaaatgagga | aattgcatcg | 3900 |
| cattgtctga | gtaggtgtca  | ttctattctg | aaaaafaaaa | tggggcagga | cagcaagggg | 3960 |
| gaggattggg | aagacaatag  | caggcatgct | ggggatgcgg | tgggctctat | gggtacccag | 4020 |
| gtgctgaaga | attgacccgg  | ttcctcctgg | gccagaaaga | agcaggcaca | tccccttctc | 4080 |
| tgtgacacac | cctgtccacg  | cccctggttc | ttagttccag | ccccactcat | aggacactca | 4140 |
| tagctcagga | gggctccgcc  | ttcaatccca | cccgctaaag | tacttggagc | ggtctctccc | 4200 |
| teceteatea | gcccaccaaa  | ccaaacctag | cctccaagag | tgggaagaaa | ttaaagcaag | 4260 |
| ataggctatt | aagtgcagag  | ggagagaaaa | tgcctccaac | atgtgaggaa | gtaatgagag | 4320 |
| aaatcataga | attttaaggc  | catgatttaa | ggccatcatg | gccttaatct | teegetteet | 4380 |
| cgctcactga | ctcgctgcgc  | teggtegtte | ggctgcggcg | agcggtatca | gctcactcaa | 4440 |
| aggcggtaat | acggttatcc  | acagaatcag | gggataacgc | aggaaagaac | atgtgagcaa | 4500 |
| aaggccagca | aaaggccagg  | aaccgtaaaa | aggccgcgtt | gctggcgttt | ttccataggc | 4560 |
| teegeeeeee | tgacgagcat  | cacaaaaatc | gacgctcaag | tcagaggtgg | cgaaacccga | 4620 |
| caggactata | aagataccag  | gcgtttcccc | ctggaagctc | cctcgtgcgc | tctcctgttc | 4680 |
| cgaccctgcc | gcttaccgga- | tacctgtccg | cctttctccc | ttcgggaagc | gtggcgcttt | 4740 |
| ctcatagctc | acgctgtagg  | tatctcagtt | cggtgtaggt | cgttcgctcc | aagctgggct | 4800 |
| gtgtgcacga | accccccgtt  | cagcccgacc | gctgcgcctt | atccggtaac | tategtettq | 4860 |
| agtccaaccc | ggtaagacac  | gacttatcgc | cactggcagc | agccactggt | aacaggatta | 4920 |
| gcagagcgag | gtatgtaggc  | ggtgctacag | agttcttgaa | gtggtggcct | aactacggct | 4980 |
| acactagaag | aacagtattt  | ggtatctgcg | ctctgctgaa | gccagttacc | ttcggaaaaa | 5040 |
| gagttggtag | ctcttgatcc  | ggcaaacaaa | ccaccgctgg | tagcggtggt | ttttttgttt | 5100 |
| gcaagcagca | gattacgcgc  | agaaaaaaag | gatctcaaga | agatcctttg | atcttttcta | 5160 |
| cggggtctga | cgctcagtgg  | aacgaaaact | cacgttaagg | gattttggtc | atgagattat | 5220 |
| caaaaaggat | cttcacctag  | atccttttaa | attaaaaatg | aagttttaaa | tcaatctaaa | 5280 |
| gtatatatga | gtaaacttgg  | tctgacagtt | accaatgctt | aatcagtgag | gcacctatct | 5340 |
| cagcgatctg | tctatttcgt  | tcatccatag | ttgcctgact | cggggggggg | gggcgctgag | 5400 |
| gtctgcctcg | tgaagaaggt  | gttgctgact | cataccagge | ctgaatcgcc | ccatcatcca | 5460 |
| gccagaaagt | gagggagcca  | cggttgatga | gagctttgtt | gtaggtggac | cagttggtga | 5520 |
| ttttgaactt | ttgctttgcc  | acggaacggt | ctgcgttgtc | gggaagatgc | gtgatctgat | 5580 |
| ccttcaactc | agcaaaagtt  | cgatttattc | aacaaagccg | ccgtcccgtc | aagtcagcgt | 5640 |
| aatgctctgc | cagtgttaca  | accaattaac | caattctgat | tagaaaaact | catcgagcat | 5700 |
| caaatgaaac | tgcaatttat  | tcatatcagg | attatcaata | ccatattttt | gaaaaagccg | 5760 |
| tttctgtaat | gaaggagaaa  | actcaccgag | gcagttccat | aggatggcaa | gatcctggta | 5820 |
| teggtetgeg | attccgactc  | gtccaacatc | aatacaacct | attaatttcc | cctcgtcaaa | 5880 |
| aataaggtta | tcaagtgaga  | aatcaccatg | agtgacgact | gaatccggtg | agaatggcaa | 5940 |
| aagcttatgc | atttctttcc  | agacttgttc | aacaggccag | ccattacgct | cgtcatcaaa | 6000 |
| atcactcgca | tcaaccaaac  | cgttattcat | tcgtgattgc | gcctgagcga | gacgaaatac | 6060 |
| gcgatcgctg | ttaaaaggac  | aattacaaac | aggaatcgaa | tgcaaccggc | gcaggaacac | 6120 |
| tgccagcgca | tcaacaatat  | tttcacctga | atcaggatat | tcttctaata | cctggaatgc | 6180 |
| tgttttcccg | gggatcgcag  | tggtgagtaa | ccatgcatca | tcaggagtac | ggataaaatg | 6240 |
| cttgatggtc | ggaagaggca  | taaattccgt | cagccagttt | agtctgacca | tctcatctgt | 6300 |
| aacatcattg | gcaacgctac  | ctttgccatg | tttcagaaac | aactctggcg | catcgggctt | 6360 |
| cccatacaat | cgatagattg  | tcgcacctga | ttgcccgaca | ttatcgcgag | cccatttata | 6420 |
| cccatataaa | tcagcatcca  | tgttggaatt | taatcgcggc | ctcgagcaag | acgtttcccg | 6480 |
|            |             |            |            |            | -          |      |

```
ttgaatatgg ctcataacac cccttgtatt actgttatg taagcagaca gttttattgt 6540 tcatgatgat atattttat cttgtgcaat gtaacatcag agattttgag acacaacgtg 6600 gctttccccc ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata 6660 catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa 6720 agtgccacct gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg 6780 tatcacgagg ccctttcgtc 6800
```

```
<210> 94
<211> 6872
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
```

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacate aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac egggacegat ecageeteeg eggeegggaa eggtgeattg gaacgeggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggetat atgecaatac tetgteette agagactgae aeggaetetg tatttttaca 1320 qgatqqqqtc ccatttatta tttacaaatt cacatataca acaacqccqt cccccqtqcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte coatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgage gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcaqtcacc qtcqtcqaca cgtgtgatca gatatcqcqq ccqctctaqa qatatcqcca 1920 ccatgcgcqt gaaggaqacc cagatgaact ggcctaacct gtqqaagtqq ggcatcctga 1980 tectgggeet ggtgateate tgeagegeea gegaeaacet gtgggtgaee gtgtactaeg 2040 gegtgeeegt gtggegegae geegaeacea eeetgttetg egeeagegae geeaaggeee 2100 acgagaccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160 cccaggaggt ggacctggag aacatcaccg agaacttcaa catgtggaag aacaacatgg 2220

|            | gcaggaggac |             |            |            |            |      |
|------------|------------|-------------|------------|------------|------------|------|
|            | cctgtgcgtg |             |            |            |            |      |
| acatcatcca | cggccccaac | atcatcggca  | acatcaccga | cgaggtgcgc | aactgcagct | 2400 |
| tcaacatgac | caccgagatc | cgcgacaaga  | agcagaaggt | gcacgccctg | ttctacaagc | 2460 |
| tggacctggt | gcagatcgag | gacaagaaca  | gcagcgagta | ccgcctgatc | aactgcaaca | 2520 |
| ccagcgtgat | caagcaggcc | tgccccaaga  | tcagcttcga | ccccatcccc | atccactact | 2580 |
| gcacccccgc | cggctacgcc | atcctgaagt  | gcaacgacaa | gaacttcaac | ggcaccggcc | 2640 |
| cctgcaccaa | cgtgagcagc | gtgcagtgca  | cccacggcat | caagcccgtg | gtgagcaccc | 2700 |
| agctgctgct | gaacggcagc | ctggccgagg  | aggagatcat | catccgcagc | gagaacctga | 2760 |
| ccaacaacgc | caagaccatc | atcgtgcacc  | tgaacaagag | cgtggagatc | aactgcaccc | 2820 |
| gccccagcaa | caacacccgc | accagcacca  | ccatcggccc | cggccaggtg | ttctaccgca | 2880 |
| ccggcgacat | caccggcaac | atccgcaagg  | cctactgcga | gatcaacggc | accaagtgga | 2940 |
| acgaggccct | gaagcaggtg | accaagaagc. | tgaaggagca | cttcaacaac | aagaccatca | 3000 |
| tettecagee | ccccagcggc | ggcgacctgg  | agatcaccat | gcaccacttc | aactgccgcg | 3060 |
| gcgagttctt | ctactgcaac | accacccgcc  | tgttcaacaa | cacctgcatc | ggcaacgaga | 3120 |
|            | ctgcaacgac |             |            |            |            |      |
|            | cgccggcctg |             |            |            |            |      |
| tgagcaacat | caccggcatc | ctgctgaccc  | gcgacggcgg | cgccaacaac | accagcaacg | 3300 |
| agatcttccg | ccccggcggc | ggcaacatca  | aggacaactg | gcgcagcgag | ctgtacaagt | 3360 |
| acaaggtggt | gcagatcgag | cccctgggca  | tcgccccac  | ccgcgccaag | cttaccgtgc | 3420 |
| aggcccgcca | gctggtgagc | ggcatcgtgc  | agcagcagag | caacctgctg | cgcgccatcg | 3480 |
| aggcccagca | gcacctgctg | cagctgaccg  | tgtggggcat | caagcagctg | caggcccgcg | 3540 |
| tgctggccgt | ggagcgctac | ctgaaggacc  | agaagttcct | cgaggagatc | tggagcaaca | 3600 |
| tgacctggat | cgagtgggag | cgcgagatca  | gcaactacac | caaccagatc | tacgagatcc | 3660 |
| tgaccgagag | ccagaaccag | caggaccgca  | acgagaagga | cctgctggag | ctggacaagt | 3720 |
| gggcgagcct | gtggagctgg | ttcgacatca  | ccaactggct | gtggtacatc | aagatettea | 3780 |
| tcatgatcgt | gggcggcctg | atcggcctgc  | gcatcatctt | cgccgtgctg | agcatctgaa | 3840 |
| cacgtgggat | ccagatctgc | tgtgccttct  | agttgccagc | catctgttgt | ttgcccctcc | 3900 |
| cccgtgcctt | ccttgaccct | ggaaggtgcc  | actcccactg | tcctttccta | ataaaatgag | 3960 |
| gaaattgcat | cgcattgtct | gagtaggtgt  | cattctattc | tggggggtgg | ggtggggcag | 4020 |
| gacagcaagg | gggaggattg | ggaagacaat  | agcaggcatg | ctggggatgc | ggtgggctct | 4080 |
| atgggtaccc | aggtgctgaa | gaattgaccc  | ggttcctcct | gggccagaaa | gaagcaggca | 4140 |
| catccccttc | tctgtgacac | accctgtcca  | cgcccctggt | tcttagttcc | agccccactc | 4200 |
| ataggacact | catagctcag | gagggctccg  | ccttcaatcc | cacccgctaa | agtacttgga | 4260 |
|            | cctccctcat |             |            |            |            |      |
|            | agataggcta |             |            |            |            |      |
|            | agaaatcata |             |            |            |            |      |
| cttccgcttc | ctcgctcact | gactcgctgc  | gctcggtcgt | tcggctgcgg | cgagcggtat | 4500 |
| cagctcactc | aaaggcggta | atacggttat  | ccacagaatc | aggggataac | gcaggaaaga | 4560 |
|            | aaaaggccag |             |            |            |            |      |
|            | gctccgcccc |             |            |            |            |      |
|            | gacaggacta |             |            |            |            |      |
|            | tccgaccctg |             |            |            |            |      |
| gcgtggcgct | ttctcatagc | tcacgctgta  | ggtatctcag | ttcggtgtag | gtcgttcgct | 4860 |
| ccaagctggg | ctgtgtgcac | gaaccccccg  | ttcagcccga | ccgctgcgcc | ttatccggta | 4920 |
|            | tgagtccaac |             |            |            |            |      |
|            | tagcagagcg |             |            |            |            |      |
|            | ctacactaga |             |            |            |            |      |
|            | aagagttggt |             |            |            |            |      |
|            | ttgcaagcag |             |            |            |            |      |
|            | tacggggtct |             |            |            |            |      |
|            | atcaaaaagg |             |            |            |            |      |
|            | aagtatatat |             |            |            |            |      |
|            | ctcagcgatc |             |            |            |            |      |
| gggggcgctg | aggtctgcct | cgtgaagaag  | gtgttgctga | ctcataccag | gcctgaatcg | 5520 |
|            |            |             |            |            |            |      |

```
ccccatcatc cagccagaaa gtgagggagc cacggttgat gagagctttg ttgtaggtgg 5580
accagttggt gattttgaac ttttgctttg ccacggaacg gtctgcgttg tcgggaagat 5640
gegtgatetg atcetteaac teageaaaag ttegatttat teaacaaage eqeeqteeeg 5700
tcaagtcagc gtaatgctct gccagtgtta caaccaatta accaattctg attagaaaaa 5760
ctcatcgagc atcaaatgaa actgcaattt attcatatca ggattatcaa taccatattt 5820
ttgaaaaagc cgtttctgta atgaaggaga aaactcaccg aggcagttcc ataggatggc 5880
aagatcctgg tatcggtctg cgattccgac tcgtccaaca tcaatacaac ctattaattt 5940
cccctcgtca aaaataaggt tatcaagtga gaaatcacca tgagtgacga ctgaatccqq 6000
tgagaatggc aaaagcttat gcatttcttt ccagacttgt tcaacaggcc agccattacq 6060
ctcgtcatca aaatcactcg catcaaccaa accgttattc attcgtgatt gcgcctgagc.6120'
gagacgaaat acgcgatcgc tgttaaaagg acaattacaa acaggaatcg aatgcaaccg 6180
gegeaggaac actgecageg cateaacaat atttteacet gaateaggat attettetaa 6240
tacctggaat gctgttttcc cggggatcgc agtggtgagt aaccatgcat catcaggagt 6300
acggataaaa tgcttgatgg tcggaagagg cataaattcc gtcagccagt ttagtctgac 6360
catchcatct glaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctqq 6420
cgcatcgggc ttcccataca atcgatagat tgtcgcacct gattgcccga cattatcgcg 6480
agcccattta tacccatata aatcagcatc catgttggaa tttaatcgcg gcctcgagca 6540
agacgtttcc cgttgaatat ggctcataac accccttgta ttactgttta tgtaagcaga 6600
cagttttatt gttcatgatg atatatttt atcttgtgca atgtaacatc agagattttg 6660
agacacaacg tggctttccc cccccccca ttattgaagc atttatcagg gttattgtct 6720
catgagegga tacatatttg aatgtattta gaaaaataaa caaatagggg tteegegeac 6780
atttccccga aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta 6840
taaaaatagg cgtatcacga ggccctttcg tc
                                                                  6872
```

<210> 95
<211> 6757
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 95

tegegegtht eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtct gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tettatgeat getatactgt ttttggettg gggeetatac accecegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200

| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260 |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | tatttttaca |      |
|            |            |            |            |            | ccccgtgcc  |      |
|            |            |            |            |            | gtgttccgga |      |
|            |            |            |            |            | ccatgcctcc |      |
|            |            |            |            |            | acttaggcac |      |
|            |            |            |            |            | gtatgtgtct |      |
|            |            |            |            |            | taaggcagcg |      |
|            |            |            |            |            | ggtaactccc |      |
|            |            |            |            |            | tgctgccgcg |      |
|            |            |            |            |            | gggtcttttc |      |
|            |            |            |            |            | gatatcgcca |      |
|            |            |            |            |            | ggtatcctgg |      |
|            |            |            |            |            | gtttactacg |      |
|            |            |            |            |            | accaaagcct |      |
|            |            |            |            |            | gacccgaacc |      |
|            |            |            |            |            | aacgacatgg |      |
|            |            |            |            |            | ccgtgcgtta |      |
|            |            |            |            |            | aacaacgtta |      |
|            |            |            |            |            | gaaatccgtg |      |
|            |            |            |            |            | ctgaaagaaa |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | tccaccatca |      |
|            |            |            |            |            | gctccggctg |      |
|            |            |            |            |            | tgcaacaacg |      |
|            |            |            |            |            | ctgctgctga |      |
|            |            |            |            |            | gacaacgtta |      |
|            |            |            |            |            | ccgaacaaca |      |
|            |            |            |            |            | ggtgacatca |      |
|            |            |            |            |            | gaaaccctga |      |
|            |            |            |            |            | aaattcgctc |      |
|            |            |            |            |            | ggtgaattct |      |
|            |            |            |            |            | accatcaccc |      |
|            |            |            |            |            | gctatgtacg |      |
|            |            |            |            |            | ctgctggttc |      |
|            |            |            |            |            | ggtaacatga |      |
|            |            |            |            |            | ccgctgggta |      |
|            |            |            |            |            | tccatcgttc |      |
|            |            |            |            |            | cagctgaccg |      |
|            |            |            |            |            | ctgaaagacc |      |
|            |            |            |            |            | atctccaact |      |
|            |            |            |            |            | aaaaacgaaa |      |
|            |            |            |            |            | atctccaact |      |
|            |            |            |            |            | gttgtttgcc |      |
|            |            |            |            |            | tcctaataaa |      |
|            |            |            |            |            | ggtggggtgg |      |
|            |            |            |            |            | gatgcggtgg |      |
|            |            |            |            |            | agaaagaagc |      |
|            |            |            |            |            | gttccagccc |      |
|            |            |            |            |            | gctaaagtac |      |
|            |            |            |            |            | ccaagagtgg |      |
|            |            |            |            |            | ctccaacatg |      |
|            |            |            |            |            | catcatggcc |      |
|            |            |            |            |            | tgcggcgagc |      |
|            |            |            |            |            | ataacgcagg |      |
| aaagaacatg | tgagcaaaag | gccagcaaaa | ggccaggaac | cgtaaaaagg | ccgcgttgct | 4500 |

```
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4560
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4620
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 4680
gggaagegtg gegetttete atageteaeg etgtaggtat eteagttegg tgtaggtegt 4740
tegetecaag etgggetgtg tgeacgaace eccegtteag eccgaceget gegeettate 4800
cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc 4860
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 4920
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 4980
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 5040
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 5100
tectttgate ttttctaegg ggtctgaege teagtggaae gaaaacteae gttaagggat 5160
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5220
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 5280
cagigaggea cetateteag egatetgiet attiegtiea tecatagitig cetgaetegg 5340
999999999 cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat accaggcctg 5400
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttgttgta 5460
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcggg 5520
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5580
tecegteaag teagegtaat getetgeeag tgttacaace aattaaceaa ttetgattag 5640
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5700
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5760
atggcaagat cctggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5820
aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5880
teeggtgaga atggeaaaag ettatgeatt tettteeaga ettgtteaac aggeeageea 5940
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 6000
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 6060
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 6120
totaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tqcatcatca 6180
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccqtcaq ccaqtttaqt 6240
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6300
tetggegeat egggetteec atacaatega tagattgteg cacetgattg eccgacatta 6360
tcgcgagccc atttataccc atataaatca gcatccatgt tggaatttaa tcgcggcctc 6420
gagcaagacg tttcccgttg aatatggctc ataacacccc ttqtattact qtttatqtaa 6480
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6540
ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta tcagggttat 6600
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6660
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6720
acctataaaa ataggcgtat cacgaggccc tttcgtc
                                                                  6757
```

```
<210> 96
<211> 6829
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 96
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgg tcagcggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
```

| +++-       |            | ~~~~++~~++ | attaaataat | +-+++-a+   | antonattao | 260  |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | aatcaattac |      |
|            |            |            |            |            | cggtaaatgg |      |
|            |            |            |            |            | cgtatgttcc |      |
|            |            |            |            |            | tacggtaaac |      |
|            |            |            |            |            | ttgacgtcaa |      |
|            |            |            |            |            | actttcctac |      |
| ttggcagtac | atctacgtat | tagtcatcgc | tattaccatg | gtgatgcggt | tttggcagta | 720  |
| catcaatggg | cgtggatagc | ggtttgactc | acggggattt | ccaagtctcc | accccattga | 780  |
|            |            |            |            |            | gtcgtaacaa |      |
|            |            |            |            |            | atataagcag |      |
|            |            |            |            |            | ttgacctcca |      |
|            |            |            |            |            | gaacgcggat |      |
|            |            |            |            |            | cccctttggc |      |
|            |            |            |            |            | ccttatgcta |      |
|            |            |            |            |            | accactcccc |      |
|            |            |            |            |            | caactatctc |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | tatttttaca |      |
|            |            |            |            |            | ccccgtgcc  |      |
|            |            |            |            | •          | gtgttccgga |      |
|            |            |            |            |            | ccatgcctcc |      |
|            |            |            |            |            | acttaggcac |      |
|            |            |            |            |            | gtatgtgtct |      |
|            |            |            |            |            | taaggcagcg |      |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg | 1800 |
|            |            |            |            |            | gggtcttttc |      |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatcgcca | 1920 |
|            |            |            |            |            | ggtatcctgg |      |
|            |            |            |            |            | gtttactacg |      |
|            |            |            |            |            | accaaagcct |      |
|            |            |            |            |            | gacccgaacc |      |
|            |            |            |            |            | aacgacatgg |      |
|            |            |            |            |            | ccgtgcgtta |      |
|            |            |            |            |            | aacaacgtta |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | gaaatccgtg |      |
|            |            |            |            |            | ctgaaagaaa |      |
|            |            |            |            |            | tccaccatca |      |
|            |            |            |            |            | gctccggctg |      |
|            |            |            |            |            | tgcaacaacg |      |
|            |            |            |            |            | ctgctgctga |      |
|            |            |            |            |            | gacaacgtta |      |
|            |            |            |            |            | ccgaacaaca |      |
|            |            |            |            |            | ggtgacatca |      |
|            |            |            |            |            | gaaaccctga |      |
| aacgtgttaa | agaaaaactg | caggaaaact | acaacaacaa | caaaaccatc | aaattcgctc | 3000 |
| cgtcctccgg | tggtgacctg | gaaatcacca | cccactcctt | caactgccgt | ggtgaattct | 3060 |
| tctactgcaa | caccacccgt | ctgttcaaca | acaacgctac | cgaagacgaa | accatcaccc | 3120 |
| tgccgtgccg | tatcaaacag | atcatcaaca | tgtggcaggg | tgttggtcgt | gctatgtacg | 3180 |
|            |            |            |            |            | ctgctggttc |      |
|            | _          |            |            |            | ggtaacatga |      |
|            |            |            |            |            | ccgctgggta |      |
|            |            |            |            |            | tccatcgttc |      |
|            |            |            |            |            | cagctgaccg |      |
|            |            |            |            |            | ctgaaagacc |      |
|            |            |            |            |            | atctccaact |      |
| ageagerega | 2466633446 |            | 22~~224462 | 2245522224 |            | 2000 |

| acaccgacac | catchaccot | ctactagaaa | actoccagae | ccagcaggaa | aaaaacgaaa               | 3660 |
|------------|------------|------------|------------|------------|--------------------------|------|
|            |            |            |            |            | atctccaact               |      |
|            |            |            |            |            | ctgcgtatca               |      |
|            |            |            |            |            | tgccagccat               |      |
|            |            |            |            |            | cccactgtcc               |      |
|            |            |            |            |            | tctattctgg               |      |
|            |            |            |            |            | aggcatgctg               |      |
|            |            |            |            |            |                          |      |
|            |            |            |            |            | tcctcctggg<br>ccctggttct |      |
|            |            |            |            |            |                          |      |
|            |            |            |            |            | tcaatcccac               |      |
|            |            |            |            |            | caaacctagc               |      |
|            |            |            |            |            | gagagaaaat               |      |
|            |            |            |            |            | atgatttaag               |      |
|            |            |            |            |            | cggtcgttcg               |      |
|            |            |            |            |            | cagaatcagg               |      |
|            |            |            |            |            | accgtaaaaa               |      |
|            |            |            |            |            | acaaaaatcg               |      |
|            |            |            |            |            | cgtttcccc                |      |
| _          |            | •          |            |            | acctgtccgc               |      |
|            |            |            |            |            | atctcagttc               |      |
|            |            |            |            |            | agcccgaccg               |      |
|            |            |            |            |            | acttatcgcc               |      |
|            |            |            |            |            | gtgctacaga               |      |
| gttcttgaag | tggtggccta | actacggcta | cactagaaga | acagtatttg | gtatctgcgc               | 5040 |
|            |            |            |            |            | gcaaacaaac               |      |
| caccgctggt | agcggtggtt | tttttgtttg | caagcagcag | attacgcgca | gaaaaaaagg               | 5160 |
| atctcaagaa | gatcctttga | tettttetae | ggggtctgac | gctcagtgga | acgaaaactc               | 5220 |
| acgttaaggg | attttggtca | tgagattatc | aaaaaggatc | ttcacctaga | tccttttaaa               | 5280 |
| ttaaaaatga | agttttaaat | caatctaaag | tatatatgag | taaacttggt | ctgacagtta               | 5340 |
| ccaatgctta | atcagtgagg | cacctatctc | agcgatctgt | ctatttcgtt | catccatagt               | 5400 |
| tgcctgactc | 999999999  | ggcgctgagg | tetgeetegt | gaagaaggtg | ttgctgactc               | 5460 |
| ataccaggcc | tgaatcgccc | catcatccag | ccagaaagtg | agggagccac | ggttgatgag               | 5520 |
| agctttgttg | taggtggacc | agttggtgat | tttgaacttt | tgctttgcca | cggaacggtc               | 5580 |
| tgcgttgtcg | ggaagatgcg | tgatctgatc | cttcaactca | gcaaaagttc | gatttattca               | 5640 |
| acaaagccgc | cgtcccgtca | agtcagcgta | atgctctgcc | agtgttacaa | ccaattaacc               | 5700 |
| aattctgatt | agaaaaactc | atcgagcatc | aaatgaaact | gcaatttatt | catatcagga               | 5760 |
| ttatcaatac | catatttttg | aaaaagccgt | ttctgtaatg | aaggagaaaa | ctcaccgagg               | 5820 |
|            |            |            |            |            | tccaacatca               |      |
|            |            |            |            |            | atcaccatga               |      |
| gtgacgactg | aatccggtga | gaatggcaaa | agcttatgca | tttctttcca | gacttgttca               | 6000 |
|            |            |            |            |            | gttattcatt               |      |
|            |            |            |            |            | attacaaaca               |      |
|            |            |            |            |            | ttcacctgaa               |      |
|            |            |            |            |            | ggtgagtaac               |      |
|            |            |            |            |            | aaattccgtc               |      |
|            |            |            |            |            | tttgccatgt               |      |
|            |            |            |            |            | cgcacctgat               |      |
|            |            |            |            |            | gttggaattt               |      |
|            |            |            |            |            | ccttgtatta               |      |
|            |            |            |            |            | ttgtgcaatg               |      |
|            |            |            |            |            | ttgaagcatt               |      |
|            |            |            |            |            | aaataaacaa               |      |
|            |            |            |            |            | aaccattatt               |      |
|            | taacctataa |            | _          |            |                          | 6829 |
|            |            |            | 555        |            |                          |      |

```
<210> 97
<211> 6752
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 97
tegegegtht eggtgatgac ggtgaaaacc tetgacacat geageteeeg gagacggtea 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcqqqtq 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac egggacegat ceagecteeg eggeegggaa eggtgeattg gaaegeggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080
tettatgeat getataetgt ttttggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagaetgae aeggaetetg tatttttaca 1320
qqatqqqqtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ecatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt teetttecat gggtetttte 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgogogt ggagggcatc cagogcaact ggaagcagtg gtggatctgg ggcatcctgg 1980
gettetggat ggtgatgate tacaacgtge geggeaacet gtgggtgace gtgtactaeg 2040
gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agetgacccc cetgtgegtg accetgeact geageaaceg ategaagaac tgeagettea 2340
acatgaccac cgaggtgcgc gacaagcgcg agaaggtgca cgccctgttc taccgcctqq 2400
acatcgtgcc cctgaagaac gagagcagca acaccagcgg cqactaccqc ctgatcaact 2460
gcaacaccag cgccatcacc caggcctgcc ccaaggtgag cttcgacccc atccccatcc 2520
actactgcgc ccccgccggc tacgccatcc tgaagtgcaa caacaagacc ttcaacqqca 2580
coggeocotg caacaacgtg agcaccatco agtgcaccca cggcaccaag cocgtggtga 2640
```

| 1          |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| gcacccagct | gctgctgaac | ggcagcctgg | ccgaggagga | gatcatcatc | cgcagcaaga | 2700 |
| acctgaccga | caacgtgaag | accatcatcg | tgcacctgaa | cgagagcgtg | gagatcaact | 2760 |
| gcacccgccc | caacaacaac | acccgcaaga | gcatccgcat | caaccccaac | caggccttct | 2820 |
| acgccaccgg | cgagatcatc | ggcgacatcc | gccaggccca | ctgcaacatc | agccgcaccg | 2880 |
| cctggaacaa | gaccctgcag | gaggtgggca | agaagctggc | cgagcacttc | cccaacaagg | 2940 |
| ccatcaagtt | cgccaagcac | agcggcggcg | acctggagat | caccacccac | agcttcaact | 3000 |
| gccgcggcga | gttcttctac | tgcaacacca | gcagcctgtt | caacagcacc | tacaccccca | 3060 |
| acagcaccga | gaacatcacc | ggcaccgaga | acagcatcat | caccatcccc | tgccgcatca | 3120 |
| agcagatcat | caacatgtgg | cagggcgtgg | gccgcgccat | gtacgccccc | cccatcgagg | 3180 |
| gcatcctgac | ctgccgcagc | aacatcaccg | gcctgctgct | gacccgcgac | ggcggcaccg | 3240 |
| gcatgcacga | caccgagatc | ttccgccccg | agggcggcga | catgcgcgac | aactggcgca | 3300 |
| gcgagctgta | caagtacaag | gtggtggaga | tcaagcccct | gggcatcgcc | cccaccaagg | 3360 |
| ccaagcttac | cgtgcaggtg | cgccagctgc | tgagcggcat | cgtgcagcag | cagagcaacc | 3420 |
| tgctgcgcgc | catcgaggcc | cagcagcaca | tgctgcagct | gaccgtgtgg | ggcatcaagc | 3480 |
| agctgcagac | ccgcgtgctg | gccatcgagc | gctacctgcg | cgaccagcag | ctcgaggaca | 3540 |
| tctggaacaa | catgacctgg | atgcagtggg | accgcgagat | cagcaactac | accaacacca | 3600 |
| tctaccgcct | gctggaggac | agccagaacc | agcaggagaa | gaacgagcag | gacctgctgg | 3660 |
| ccctggacaa | gtggcagaac | ctgtggacct | ggttcggcat | caccaactgg | ctgtggtgaa | 3720 |
| cacgtgggat | ccagatctgc | tgtgccttct | agttgccagc | catctgttgt | ttgcccctcc | 3780 |
| cccgtgcctt | ccttgaccct | ggaaggtgcc | actcccactg | tcctttccta | ataaaatgag | 3840 |
| gaaattgcat | cgcattgtct | gagtaggtgt | cattctattc | tggggggtgg | ggtggggcag | 3900 |
| gacagcaagg | gggaggattg | ggaagacaat | agcaggcatg | ctggggatgc | ggtgggctct | 3960 |
| atgggtaccc | aggtgctgaa | gaattgaccc | ggttcctcct | gggccagaaa | gaagcaggca | 4020 |
| catccccttc | tctgtgacac | accctgtcca | cgcccctggt | tcttagttcc | agccccactc | 4080 |
| ataggacact | catagctcag | gagggctccg | ccttcaatcc | cacccgctaa | agtacttgga | 4140 |
| gcggtctctc | cctccctcat | cagcccacca | aaccaaacct | agcctccaag | agtgggaaga | 4200 |
| aattaaagca | agataggcta | ttaagtgcag | agggagagaa | aatgcctcca | acatgtgagg | 4260 |
| aagtaatgag | agaaatcata | gaattttaag | gccatgattt | aaggccatca | tggccttaat | 4320 |
| cttccgcttc | ctcgctcact | gactcgctgc | gctcggtcgt | teggetgegg | cgagcggtat | 4380 |
| cageteacte | aaaggcggta | atacggttat | ccacagaatc | aggggataac | gcaggaaaga | 4440 |
| acatgtgagc | aaaaggccag | caaaaggcca | ggaaccgtaa | aaaggccgcg | ttgctggcgt | 4500 |
| ttttccatag | gctccgcccc | cctgacgagc | atcacaaaaa | tcgacgctca | agtcagaggt | 4560 |
| ggcgaaaccc | gacaggacta | taaagatacc | aggcgtttcc | ccctggaagc | tecetegtge | 4620 |
| gctctcctgt | tccgaccctg | ccgcttaccg | gatacctgtc | cgcctttctc | ccttcgggaa | 4680 |
| gcgtggcgct | ttctcatagc | tcacgctgta | ggtatctcag | ttcggtgtag | gtcgttcgct | 4740 |
| ccaagctggg | ctgtgtgcac | gaaccccccg | ttcagcccga | ccgctgcgcc | ttatccggta | 4800 |
| actatcgtct | tgagtccaac | ccggtaagac | acgacttatc | gccactggca | gcagccactg | 4860 |
| gtaacaggat | tagcagagcg | aggtatgtag | gcggtgctac | agagttcttg | aagtggtggc | 4920 |
| ctaactacgg | ctacactaga | agaacagtat | ttggtatctg | cgctctgctg | aagccagtta | 4980 |
| ccttcggaaa | aagagttggt | agctcttgat | ccggcaaaca | aaccaccgct | ggtagcggtg | 5040 |
| gtttttttgt | ttgcaagcag | cagattacgc | gcagaaaaaa | aggatctcaa | gaagatcctt | 5100 |
| tgatcttttc | tacggggtct | gacgctcagt | ggaacgaaaa | ctcacgttaa | gggattttgg | 5160 |
| tcatgagatt | atcaaaaagg | atcttcacct | agatcctttt | aaattaaaaa | tgaagtttta | 5220 |
| aatcaatcta | aagtatatat | gagtaaactt | ggtctgacag | ttaccaatgc | ttaatcagtg | 5280 |
| aggcacctat | ctcagcgatc | tgtctatttc | gttcatccat | agttgcctga | ctcggggggg | 5340 |
| gggggcgctg | aggtctgcct | cgtgaagaag | gtgttgctga | ctcataccag | gcctgaatcg | 5400 |
| ccccatcatc | cagccagaaa | gtgagggagc | cacggttgat | gagagetttg | ttgtaggtgg | 5460 |
| accagttggt | gattttgaac | ttttgctttg | ccacggaacg | gtctgcgttg | tcqqqaaqat | 5520 |
| gcgtgatctg | atcettcaaç | tcagcaaaag | ttcgatttat | tcaacaaagc | caccatecea | 5580 |
| tcaagtcagc | gtaatgctct | gccagtgtta | caaccaatta | accaattctg | attagaaaaa | 5640 |
| ctcatcgage | atcaaatgaa | actgcaattt | attcatatca | ggattatcaa | taccatatth | 5700 |
| ttgaaaaagc | cgtttctqta | atgaaggaga | aaactcacco | aggcagttcc | ataggatgor | 5760 |
| aagatcctqq | tateggtetg | cgattccqac | tcqtccaaca | tcaatacaac | ctattaatth | 5820 |
| cccctcgtca | aaaataaggt | tatcaaqtqa | qaaatcacca | tgagtgacga | ctgaatccoo | 5880 |
| tgagaatqqc | aaaagcttat | gcatttcttt | ccagacttot | tcaacaggcc | agccattaco | 5940 |
|            | J. 2       |            |            |            | 2-2-000    |      |

```
ctcgtcatca aaatcactcg catcaaccaa accgttattc attcgtgatt gcgcctgagc 6000
gagacgaaat acgcgatcgc tgttaaaagg acaattacaa acaggaatcg aatgcaaccg 6060
gegeaggaac actgecageg cateaacaat atttteacet gaateaggat attettetaa 6120
tacctggaat gctgttttcc cggggatcgc agtggtgagt aaccatgcat catcaggagt 6180
acggataaaa tgcttgatgg tcggaagagg cataaattcc gtcagccagt ttagtctgac 6240
catctcatct gtaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctgg 6300
egeateggge tteccataca ategatagat tgtegeacet gattgeeega cattategeg 6360
ageceattta tacceatata aateageate eatgttggaa tttaategeg geetegagea 6420
agacgtttcc cgttgaatat ggctcataac accccttgta ttactgttta tgtaagcaga 6480
cagtittatt gticatgatg atatatittt atcitgtgca atgtaacatc agagatittg 6540
agacacaacg tggctttccc cccccccca ttattgaagc atttatcagg gttattqtct 6600
catgagegga tacatatttg aatgtattta gaaaaataaa caaatagggg tteegegeac 6660
atttccccga aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta 6720
taaaaatagg cgtatcacga ggccctttcg tc
<210> 98
<211> 6623
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegettt eggegatgae ggegaaaace teegacacat geageteeeg gagaeggea 60
cagettgtct gtaageggat geegggagea gacaageeeg teaggggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgq 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cqqtaaatqq 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agelegitta gigaacegic agalegeetg gagaegeeal ceaegetgit tigaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
```

catgggetet teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 ageacaatge ceaceaceae eagtgtgeeg eacaaggeeg tggeggtagg gtatgtgtet 1620 gaaaatgage gtggagattg ggetegeaeg getgaegeag atggaagaet taaggeageg 1680 geagaagaag atgeaggeag etgagttgt gtattetgat aagagteaga ggtaacteee 1740

| gttgcggtgc | tgttaacggt | ggagggcagt  | gtagtctgag | cagtactcgt | tgctgccgcg  | 1800 |
|------------|------------|-------------|------------|------------|-------------|------|
| cgcgccacca | gacataatag | ctgacagact  | aacagactgt | tectttccat | gggtcttttc  | 1860 |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca  | gatatcgcgg | ccgctctaga | gatatcgcca  | 1920 |
|            | ggagggcatc |             |            |            |             |      |
| gcttctggat | ggtgatgatc | tacaacgtgc  | gcggcaacct | gtgggtgacc | gtqtactacq  | 2040 |
| gcgtgcccgt | gtggaaggag | gccaagacca  | ccctqttctq | cqccaqcqac | gccaaggcct  | 2100 |
| acgacgccga | ggtgcacaac | gtgtgggcca  | cccacqcctq | catacccacc | gaccccaacc  | 2160 |
| cccaggagat | ggtgctggag | aacqtqaccq  | agaacttcaa | catgtgggag | aacqacatqq  | 2220 |
| tggagcagat | gcaccaggac | atcatcagcc  | tataggacca | gageetgaag | ccctacataa  | 2280 |
| agctgacccc | cctgtgcgtg | accetquact  | gcagcaaccg | atcgatcaac | tgcaacacca  | 2340 |
|            | ccaggcctgc |             |            |            |             |      |
|            | ctacgccatc |             |            |            |             |      |
| qcaacaacqt | gagcaccatc | cagtgcaccc  | acggcaccaa | acccataata | agcacccagc  | 2520 |
|            | cggcagcctg |             |            |            |             |      |
|            | gaccatcatc |             |            |            |             |      |
| ccaacaacaa | cacccgcaag | agcaticcgca | tcaaccccaa | ccagacette | tacaccacca  | 2700 |
| gcgagatcat | cggcgacatc | caccadaccc  | actocaacat | caggeetee  | acctageacca | 2760 |
| agaccctgca | ggaggtgggc | aagaagetgg  | ccaaccactt | ccccaacaac | gcccggaaca  | 2020 |
|            | cagcggcggc |             |            |            |             |      |
|            | ctgcaacacc |             |            |            |             |      |
| agaacatcac | cggcaccgag | ageageetge  | teaccatece | ctacaccccc | aacagcaccg  | 2940 |
|            |            |             |            |            |             |      |
| cctaccacaa | gcagggcgtg | ggccgcgcca  | taaccccca  | ccccatcgag | ggcateetga  | 3120 |
| acaccaacaa | caacatcacc | ggcccgccgc  | agacccgcga | cggcggcacc | ggcatgcatg  | 3120 |
| acaccgagac | cttccgcccc | atoaaggeg   | togggatage | caactggege | agegagetgt  | 3180 |
| acaagcacaa | ggtggtggag | atcaagcccc  | tagtage    | ccccaccaag | gccaagccca  | 3240 |
| ccgtgcaggt | gcgccagctg | ctgageggea  | tegegeagea | gcagagcaac | cracracaca  | 3300 |
| ccaccyagge | ccagcagcac | acgetgeage  | cgacegtgtg | gggcatcaag | cagctgcaga  | 3360 |
| acataacata | ggccatcgag | cgctacetge  | gegaecagea | getegaggae | atctggaaca  | 3420 |
| tastassass | gatgcagtgg | gaccgcgaga  | ccagcaacta | Caccaacacc | acctaccgcc  | 3480 |
| cgccggagga | cagccagaac | cagcaggaga  | agaacgagca | ggacetgetg | gccctggaca  | 3540 |
| agrygeagaa | cctgtggacc | tggtteggea  | teaceaactg | getgtggtga | acacgtggga  | 3600 |
| tecagatetg | ctgtgccttc | tagttgccag  | ccatctgttg | tttgcccctc | ccccgtgcct  | 3660 |
| teettgaeee | tggaaggtgc | cactcccact  | greerreer  | aataaaatga | ggaaattgca  | 3720 |
| tegeattgte | tgagtaggtg | tcattctatt  | ctggggggtg | gggtggggca | ggacagcaag  | 3780 |
| ggggaggacc | gggaagacaa | tagcaggcat  | gctggggatg | cggtgggctc | tatgggtacc  | 3840 |
| caggtgctga | agaattgacc | cggttcctcc  | tgggccagaa | agaagcaggc | acateceett  | 3900 |
| ctctgtgaca | caccctgtcc | acgcccctgg  | ttcttagttc | cagccccact | cataggacac  | 3960 |
| tcatagetea | ggagggctcc | gccttcaatc  | ccacccgcta | aagtacttgg | ageggtetet  | 4020 |
| ccctccctca | tcagcccacc | aaaccaaacc  | tagcctccaa | gagtgggaag | aaattaaagc  | 4080 |
|            | attaagtgca |             |            |            |             |      |
| gagaaatcat | agaattttaa | ggccatgatt  | taaggccatc | atggccttaa | tetteegett  | 4200 |
| cctcgctcac | tgactcgctg | cgctcggtcg  | ttcggctgcg | gcgagcggta | tcagctcact  | 4260 |
| caaaggcggt | aatacggtta | tccacagaat  | caggggataa | cgcaggaaag | aacatgtgag  | 4320 |
|            | gcaaaaggcc |             |            |            |             |      |
|            | ccctgacgag |             |            |            |             |      |
| cgacaggact | ataaagatac | caggcgtttc  | cccctggaag | ctccctcgtg | cgctctcctg  | 4500 |
| ttccgaccct | gccgcttacc | ggatacctgt  | ccgccttct  | cccttcggga | agcgtggcgc  | 4560 |
| tttctcatag | ctcacgctgt | aggtatctca  | gttcggtgta | ggtcgttcgc | tccaagctgg  | 4620 |
| gctgtgtgca | cgaacccccc | gttcagcccg  | accgctgcgc | cttatccggt | aactatcgtc  | 4680 |
|            | cccggtaaga |             |            |            |             |      |
|            | gaggtatgta |             |            |            |             |      |
| gctacactag | aagaacagta | tttggtatct  | gcgctctgct | gaagccagtt | accttcggaa  | 4860 |
| aaagagttgg | tagctcttga | tccggcaaac  | aaaccaccgc | tggtagcggt | ggtttttta   | 4920 |
| tttgcaagca | gcagattacg | cgcagaaaaa  | aaggatctca | agaagateet | ttgatctttt  | 4980 |
| ctacggggtc | tgacgctcag | tggaacgaaa  | actcacgtta | agggattttq | gtcatqaqat  | 5040 |
|            |            |             | -          |            |             |      |

```
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 5100
aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta 5160
teteagegat etgtetattt egtteateea tagttgeetg aetegggggg ggggggeget 5220
gaggtctgcc tcgtgaagaa ggtgttgctg actcatacca ggcctgaatc gccccatcat 5280
ccagccagaa agtgagggag ccacggttga tgagagcttt gttgtaggtg gaccagttgg 5340
tgattttgaa cttttgcttt gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct 5400
gatecticaa eteageaaaa gitegatita tieaaeaaag eegeegteee gicaagteag 5460
cgtaatgctc tgccagtgtt acaaccaatt aaccaattct gattagaaaa actcatcgag 5520
catcaaatga aactgcaatt tattcatatc aggattatca ataccatatt tttgaaaaag 5580
ccgtttctgt aatgaaggag aaaactcacc gaggcagttc cataggatgg caagatcctg 5640
gtateggtet gegatteega etegteeaac ateaatacaa eetattaatt teecetegte 5700
aaaaataagg ttatcaagtg agaaatcacc atgagtgacg actgaatccg gtgagaatgg 5760
caaaagetta tgcatttett tecagaettg tteaacagge cagecattae getegteate 5820
aaaatcactc gcatcaacca aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa 5880
tacgcgatcg ctgttaaaag gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa 5940
cactgccagc gcatcaacaa tattttcacc tgaatcagga tattcttcta atacctggaa 6000
tgctgttttc ccggggatcg cagtggtgag taaccatgca tcatcaggag tacggataaa 6060
atgcttgatg gtcggaagag gcataaattc cgtcagccag tttagtctga ccatctcatc 6120
tgtaacatca ttggcaacgc tacctttgcc atgtttcaga aacaactctg gcgcatcggg 6180
cttcccatac aatcgataga ttgtcgcacc tgattgcccg acattatcgc gagcccattt 6240
atacccatat aaatcagcat ccatgttgga atttaatcgc ggcctcgagc aagacgtttc 6300
ccgttgaata tggctcataa caccccttgt attactgttt atgtaagcag acagttttat 6360
tgttcatgat gatatatttt tatcttgtgc aatgtaacat cagagatttt gagacacaac 6420
gtggetttee ecceecec attattgaag catttateag ggttattgte teatgagegg 6480
atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg 6540
aaaagtgcca cctgacgtct aagaaaccat tattatcatg acattaacct ataaaaataq 6600
gcgtatcacg aggccctttc gtc
                                                                  6623
```

```
<211> 6542
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 99
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggbcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataaqcaq 900
agetegttta gtgaaccgtc agategeetg gagacgecat ceacqetgtt ttgaccteca 960
```

<210> 99

| tagaagacac | coggaccgat   | ccagectecg   | cggccgggaa   | cggtgcattg   | gaacgcggat   | 1020 |
|------------|--------------|--------------|--------------|--------------|--------------|------|
| teceratace | aagagtgacg   | taaqtaccqc   | ctatagactc   | tataggcaca   | ccctttggc    | 1080 |
| tettateeat | actatactat   | ttttaactta   | gggcctatac   | acccccgctt   | ccttatgcta   | 1140 |
| tanntnatnn | tatagettag   | cctataggtg   | tagattatta   | accattattg   | accactcccc   | 1200 |
| tattootoac | gatactttcc   | attactaatc   | cataacatgg   | ctctttgcca   | caactatctc   | 1260 |
| tattggtgat | atoccaatac   | tetateette   | agagactgac   | acqqactctq   | tatttttaca   | 1320 |
| cattggctat | ccatttatta   | tttacaaatt   | cacatataca   | acaacgccqt   | ccccgtgcc    | 1380 |
| ggarggggcc | attaaacata   | acatagaata   | trcacacaaa   | tctcqqqtac   | gtgttccgga   | 1440 |
| cgcagucuu  | tatacactaa   | conconanct   | tecacateca   | agccctggtc   | ccatgcctcc   | 1500 |
| catgggetet | catacatag    | cagergagee   | ctcctaacag   | tagaggccag   | acttaggcac   | 1560 |
| ageggeteat | ggccgcccgg   | cageteeteg   | cacaadacca   | taacaataaa   | gtatgtgtct   | 1620 |
| ageacaatge | ccaccaccac   | cagtgagacg   | actaecacea   | atggaagact   | taaggcagcg   | 1680 |
| gaaaacgagc | gragagagag   | ctaaattatt   | graffctgat   | aagagtcaga   | ggtaactccc   | 1740 |
| gcagaagaag | acgcaggcag   | canagacaat   | gtactetgae   | cagtactcgt   | tgctgccgcg   | 1800 |
| getgeggtge | cyccaacyyc   | ggagggcagt   | aacacactot   | teettteeat   | gggtcttttc   | 1860 |
| cgcgccacca | gacataatag   | ccgacagacc   | datatogo     | ccactctaga   | gatatogcca   | 1920 |
| tgcagtcacc | geegeegaca   | egegegacea   | gacaccacta   | atacatetaa   | gatategeea   | 1980 |
| ccatgcgcgt | ggagggcatc   | cagegeaact   | ggaagcagcg   | gtggatetgg   | ggcatcctgg   | 2040 |
| gettetggat | ggtgatgate   | tacaacgtgc   | geggeaacee   | gegggegaee   | gtgtactacg   | 2100 |
| gegtgeeegt | grggaaggag   | gccaagacca   | eccigitety   | cgccagcgac   | gccaaggcct   | 2160 |
| acgacgccga | ggtgcacaac   | grgrgggcca   | cecaegeerg   | cgtgtttaaaa  | gaccccaacc   | 2220 |
| cccaggagat | ggtgctggag   | aacgtgaccg   | agaacttcaa   | catgtgggag   | aacgacatgg   | 2280 |
| tggagcagat | gcaccaggac   | arcarcagec   | tgtgggacea   | gageetgaag   | ccctgcgtga   | 2340 |
| agctgacccc | cctgtgcgtg   | accetgeact   | gcagcaaccg   | accyatcaac   | tgcaacacca   | 2340 |
| gcgccatcac | ccaggcctgc   | cccaaggtga   | gcttcgaccc   | catececate   | cactactgcg   | 2460 |
| ccccgccgg  | ctacgccatc   | ctgaagtgca   | acaacaagac   | cttcaacggc   | accggcccct   | 2400 |
| gcaacaacgt | gagcaccatc   | cagtgcaccc   | acggcaccaa   | gcccgtggtg   | agcacccagc   | 2520 |
| tgctgctgaa | cggcagcctg   | gccgaggagg   | agatcatcat   | ccgcagcaag   | aacctgaccg   | 2580 |
| acaacgtgaa | gaccatcatc   | gtgcacctga   | acgagagcgt   | ggagatcaac   | tgcacccgcc   | 2640 |
| gatcggccca | ctgcaacatc   | agccgcaccg   | cctggaacaa   | gaccctgcag   | gaggtgggca   | 2700 |
| agaagctggc | cgagcacttc   | cccaacaagg   | ccatcaagtt   | cgccaagcac   | agcggcggcg   | 2760 |
| acctggagat | caccacccac   | agcttcaact   | gccgcggcga   | gttettetae   | tgcaacacca   | 2820 |
| gcagcctgtt | caacagcacc   | tacaccccca   | acagcaccga   | gaacatcacc   | ggcaccgaga   | 2880 |
| acagcatcat | caccatcccc   | tgccgcatca   | agcagatcat   | caacatgtgg   | cagggcgtgg   | 2940 |
| gccgcgccat | gtacgccccc   | cccatcgagg   | gcatcctgac   | ctgccgcagc   | aacatcaccg   | 3000 |
| gcctgctgct | gacccgcgac   | ggcggcaccg   | gcatgcacga   | caccgagate   | ttccgccccg   | 3060 |
| agggcggcga | catgcgcgac   | : aactggcgca | gcgagctgta   | caagtacaag   | gtggtggaga   | 3120 |
| tcaaqcccct | gggcatcgcc   | cccaccaagg   | ccaagcttac   | cgtgcaggtg   | cgccagctgc   | 3180 |
| taaacaacat | cqtqcaqcaq   | cagagcaacc   | tgctgcgcgc   | catcgaggco   | cagcagcaca   | 3240 |
| tactacaact | gaccgtgtgg   | ggcatcaago   | agctgcagac   | ccgcgtgctg   | gccatcgagc   | 3300 |
| gctacctgcg | cgaccagcag   | ctcgaggaca   | . tctggaacaa | catgacctgg   | , atgcagtggg | 3360 |
| accacaaaat | caqcaactac   | accaacacca   | tctaccgcct   | gctggaggac   | : agccagaacc | 3420 |
| agcaggagaa | gaacgagcag   | gacctgctgg   | ccctggacaa   | . gtggcagaac | : ctgtggacct | 3480 |
| gattcagcat | caccaactgo   | ı ctgtggtgaa | ı cacgtgggat | ccagatetgo   | : tgtgccttct | 3540 |
| agttqccaqc | : catctgttgt | ttgcccctcc   | : cccgtgcctt | ccttgaccct   | : ggaaggtgcc | 3600 |
| actoccacto | r tootttoota | ataaaatgag   | , gaaattgcat | cgcattgtct : | : gagtaggtgt | 3660 |
| cattetatte | : taaaaaatac | , qqtqqqqcaq | gacagcaagg   | , gggaggattg | g ggaagacaat | 3720 |
| agcaggcato | r ctggggatgd | : ggtgggctct | : atgggtacco | : aggtgctgaa | a gaattgaccc | 3/80 |
| aattectect | gggccagaaa   | a gaagcaggca | a catcccctto | : tctgtgacac | : accctgtcca | 3840 |
| cacccctaat | : tcttagttcc | : agccccacto | : ataggacact | : catageteag | g gagggctccg | 3900 |
| ccttcaatco | cacccgctaa   | a agtacttgga | a geggtetete | : cctccctcat | c cagcccacca | 3960 |
| aaccaaacct | : agcctccaac | agtgggaaga   | a aattaaagca | agataggcta   | a ttaagtgcag | 4020 |
| aggagagaa  | aatgcctcc    | acatgtgagg   | g aagtaatgag | g agaaatcata | a gaattttaag | 4080 |
| accataatti | aaggccatca   | a togccttaat | cttccgcttc   | ctcgctcact   | gactcgctgc   | 4140 |
| actcaatcat | teggetaca    | g cqaqcqqtat | cagetcacte   | aaaggcggt:   | a atacggttat | 4200 |
| ccacadaato | agggataa     | gcaggaaag    | a acatgtgage | aaaaggccag   | g caaaaggcca | 4260 |
|            |              | 5 55 5       | J-2-2        |              |              |      |

```
ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgccc cctgacgagc 4320
atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc 4380
aggogithed deciggaage teectogige geteteetgt teegacootig cogettaceg 4440
gatacetgte egeetttete eettegggaa gegtggeget tteteatage teaegetgta 4500
ggtateteag tteggtgtag gtegtteget ccaagetggg etgtgtgcae gaaceeeeeg 4560
ttcagcccga ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggtaagac 4620
acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag 4680
geggtgetac agagttettg aagtggtgge etaactaegg etacactaga agaacagtat 4740
ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt agctcttgat 4800
ccggcaaaca aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag cagattacgc 4860
gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgctcagt 4920
ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct 4980
agateetttt aaattaaaaa tgaagtttta aateaateta aagtatatat gagtaaaett 5040
ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc 5100
gttcatccat agttgcctga ctcggggggg gggggcgctg aggtctgcct cgtgaagaag 5160
gtgttgctga ctcataccag gcctgaatcg ccccatcatc cagccagaaa gtgagggagc 5220
cacggttgat gagagetttg ttgtaggtgg accagttggt gattttgaac ttttgctttg 5280
ccaeggaacg gtctgcgttg tcgggaagat gcgtgatctg atccttcaac tcagcaaaag 5340
ttegathtat teaacaaage egeegteeeg teaagteage gtaatgetet geeagtgtta 5400
caaccaatta accaattctq attaqaaaaa ctcatcqaqc atcaaatqaa actgcaattt 5460
attcatatca ggattatcaa taccatattt ttgaaaaagc cgtttctgta atgaaggaga 5520
aaactcaccg aggcagttcc ataggatggc aagatcctgg tatcggtctg cgattccgac 5580
tcgtccaaca tcaatacaac ctattaattt cccctcgtca aaaataaggt tatcaagtga 5640
gaaatcacca tgagtgacga ctgaatccgg tgagaatggc aaaagcttat gcatttcttt 5700
ccagacttgt tcaacaggcc agccattacg ctcgtcatca aaatcactcg catcaaccaa 5760
acceptatte attegtgatt gegeetgage gagacgaaat acgegatege tgttaaaagg 5820
acaattacaa acaggaatcg aatgcaaccg gcgcaggaac actgccagcg catcaacaat 5880
attttcacct gaatcaggat attcttctaa tacctggaat gctgttttcc cggggatcgc 5940
agtggtgagt aaccatgcat catcaggagt acggataaaa tgcttgatgg tcggaagagg 6000
acctttgcca tgtttcagaa acaactctgg cgcatcgggc ttcccataca atcgatagat 6120
tgtcgcacct gattgcccga cattatcgcg agcccattta tacccatata aatcagcatc 6180
catgttggaa tttaatcgcg gcctcgagca agacgtttcc cgttgaatat ggctcataac 6240
accccttgta ttactgttta tgtaagcaga cagttttatt gttcatgatg atatattttt 6300
atcttgtgca atgtaacatc agagattttg agacacaacg tggctttccc cccccccca 6360
ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta 6420
gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta 6480
agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg 6540
tc
                                                                6542
```

```
<210> 100
<211> 6473
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 100
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgg tcagcggttg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
```

|            |            |            |            |            | ttggctcatg |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | aatcaattac |      |
|            |            |            |            |            | cggtaaatgg |      |
|            |            |            |            |            | cgtatgttcc |      |
|            |            |            |            |            | tacggtaaac |      |
|            |            |            |            |            | ttgacgtcaa |      |
|            |            |            |            |            | actttcctac |      |
| ttggcagtac | atctacgtat | tagtcatcgc | tattaccatg | gtgatgcggt | tttggcagta | 720  |
| catcaatggg | cgtggatagc | ggtttgactc | acggggattt | ccaagtctcc | accccattga | 780  |
|            |            |            |            |            | gtcgtaacaa |      |
|            |            |            |            |            | atataagcag |      |
| agctcgttta | gtgaaccgtc | agatcgcctg | gagacgccat | ccacgctgtt | ttgacctcca | 960  |
|            |            |            |            |            | gaacgcggat |      |
| tccccgtgcc | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca | cccctttggc | 1080 |
| tcttatgcat | gctatactgt | ttttggcttg | gggcctatac | acccccgctt | ccttatgcta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg | tgggttattg | accattattg | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260 |
| tattggctat | atgccaatac | tetgteette | agagactgac | acggactctg | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacgccgt | ccccgtgcc  | 1380 |
| cgcagttttt | attaaacata | gcgtgggatc | tccacgcgaa | tetegggtae | gtgttccgga | 1440 |
| catgggctct | tctccggtag | cggcggagct | tccacatccg | agccctggtc | ccatgcctcc | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagaçt | taaggcagcg | 1680 |
|            |            |            |            |            | ggtaactccc |      |
|            |            |            |            |            | tgctgccgcg |      |
|            |            |            |            |            | gggtcttttc |      |
|            |            |            |            |            | gatatcgcca |      |
|            |            |            |            |            | ggcatcctgg |      |
|            |            |            |            |            | gtgtactacg |      |
|            |            |            |            |            | gccaaggcct |      |
|            |            |            |            |            | gaccccaacc |      |
|            |            |            |            |            | aacgacatgg |      |
|            |            |            |            |            | ccctgcgtga |      |
|            |            |            |            |            | tgcaacacca |      |
|            |            |            |            |            | cactactgcg |      |
|            |            |            |            |            | accggcccct |      |
|            |            |            |            |            | agcacccagc |      |
|            |            |            |            |            | aacctgaccg |      |
|            |            |            |            |            | tgcacccgcc |      |
|            |            |            |            |            | gaggtgggca |      |
|            |            |            |            |            | agcggcggcg |      |
|            |            |            |            |            | tgcaacaccc |      |
|            |            |            |            |            | ggccgcgcca |      |
|            |            |            |            |            | ggcctgctgc |      |
|            |            |            |            |            | gaggggggg  |      |
|            |            |            |            |            | atcaagcccc |      |
|            |            |            |            |            | ctgagcggca |      |
|            |            |            |            |            | atgctgcagc |      |
|            |            |            |            |            | cgctacctgc |      |
|            |            |            |            |            | gaccgcgaga |      |
|            |            |            |            |            | cagcaggaga |      |
|            |            |            |            |            | tggttcggca |      |
|            |            |            |            |            | tagttgccag |      |
|            |            |            |            |            | cactcccact |      |
|            |            |            | -          |            |            |      |

| gtcctttcct | aataaaatga | ggaaattgca | togcattgto | tgagtaggtg | tcattctatt | 3600 |
|------------|------------|------------|------------|------------|------------|------|
|            | gggtggggca |            |            |            |            |      |
| gctggggatg | cggtgggctc | tatgggtacc | caggtgctga | agaattgacc | cggttcctcc | 3720 |
|            | agaagcaggc |            |            |            |            |      |
|            | cagccccact |            |            |            |            |      |
|            | aagtacttgg |            |            |            |            |      |
|            | gagtgggaag |            |            |            |            |      |
|            | aacatgtgag |            |            |            |            |      |
|            | atggccttaa |            |            |            |            |      |
|            | gcgagcggta |            |            |            |            |      |
|            | cgcaggaaag |            |            |            |            |      |
|            | gttgctggcg |            |            |            |            |      |
|            | aagtcagagg |            |            |            |            |      |
| cccctggaag | ctccctcgtg | cactctccta | ttccqaccct | qccqcttacc | ggatacctgt | 4380 |
|            | cccttcggga |            |            |            |            |      |
| gttcggtgta | ggtcgttcgc | tccaagctgg | gctgtgtgca | cqaacccccc | gttcagcccg | 4500 |
| accgctgcgc | cttatccggt | aactatcqtc | ttgagtccaa | cccqqtaaqa | cacgacttat | 4560 |
|            | agcagccact |            |            |            |            |      |
|            | gaagtggtgg |            |            |            |            |      |
|            | gaagccagtt |            |            |            |            |      |
| aaaccaccgc | tggtagcggt | gatttttta  | tttgcaagca | gcagattacg | cacagaaaaa | 4800 |
| aaggatetea | agaagatcct | ttgatctttt | ctacqqqqtc | tgacgctcag | tggaacgaaa | 4860 |
| actcacqtta | agggatttg  | gtcatgagat | tatcaaaaag | gatcttcacc | tagatecttt | 4920 |
| taaattaaaa | atgaagtttt | aaatcaatct | aaagtatata | tgagtaaact | tagtetaaca | 4980 |
| gttaccaatq | cttaatcagt | gaggcaccta | tctcagcgat | ctgtctattt | cattcatcca | 5040 |
| tagttgcctg | actcgggggg | gagagacact | gaggtetgee | tcataaaaaa | gatattacta | 5100 |
| actcatacca | ggcctgaatc | gccccatcat | ccagccagaa | agtgaggag  | ccacaattaa | 5160 |
| tgagagettt | gttgtaggtg | gaccagttgg | tgattttgaa | cttttacttt | gccacggaac | 5220 |
| ggtctgcgtt | gtcgggaaga | tacataatct | gatcetteaa | ctcagcaaaa | gttcgattta | 5280 |
| ttcaacaaaq | ccgccgtccc | gtcaagtcag | cataatactc | taccagtatt | acaaccaatt | 5340 |
| aaccaattct | gattagaaaa | actcatcgag | catcaaatga | aactgcaatt | tattcatatc | 5400 |
| aggattatca | ataccatatt | tttgaaaaag | ccatttctat | aatgaaggag | aaaactcacc | 5460 |
| gaggcagttc | cataggatgg | caagatcctq | gtateggtet | gcgattccga | ctcqtccaac | 5520 |
| atcaatacaa | cctattaatt | tececteate | aaaaataaqq | ttatcaagtg | agaaatcacc | 5580 |
|            | actgaatccg |            |            |            |            |      |
| ttcaacaggc | cagccattac | gctcgtcatc | aaaatcactc | gcatcaacca | aaccgttatt | 5700 |
| cattcgtgat | tgcgcctgag | cgagacgaaa | tacgcgatcg | ctqttaaaaq | gacaattaca | 5760 |
|            | gaatgcaacc |            |            |            |            |      |
| tgaatcagga | tattcttcta | atacctggaa | tgctgttttc | ccqqqqatcq | cagtggtgag | 5880 |
| taaccatgca | tcatcaggag | tacqqataaa | atgettgatg | qtcqqaaqaq | qcataaattc | 5940 |
| cgtcagccag | tttagtctga | ccatctcatc | tgtaacatca | ttggcaacgc | tacctttqcc | 6000 |
|            | aacaactctg |            |            |            |            |      |
|            | acattatege |            |            |            |            |      |
| atttaatcgc | ggcctcgagc | aagacgtttc | ccgttgaata | tggctcataa | caccccttat | 6180 |
|            | atgtaagcag |            |            |            |            |      |
|            | cagagatttt |            |            |            |            |      |
|            | ggttattgtc |            |            |            |            |      |
|            | gttccgcgca |            |            |            |            |      |
|            | acattaacct |            |            |            |            | 6473 |
|            |            |            |            |            | -          | -    |

<210> 101

<211> 6554

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 101 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatqttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tettatgeat getatactgt tittggettg gggeetatac acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980 gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040 gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100 acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160 cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220 tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agetgacece cetgtgegtg accetgeact geageaaceg ategateaac tgcaacacca 2340 gegecateae ecaggeetge eccaaggtga gettegacee catececate cactactgeg 2400 ecceegeegg ctaegecate etgaagtgea acaacaagae etteaaegge accggeecet 2460 gcaacaacgt gagcaccatc cagtgcaccc acggcaccaa gcccgtggtg agcacccagc 2520 tgctgctgaa cggcagcctg gccgaggagg agatcatcat ccgcagcaag aacctgaccg 2580 acaacgtgaa gaccatcatc gtgcacctga acgagagcgt ggagatcaac tgcacccgcc 2640 ccaacaacaa caccegcaag agcateegca teggeecegg ccaggeette tacgecaceg 2700 gcgagatcat cggcgacatc cgccaggccc actgcaacat cagccgcacc gcctggaaca 2760 agaccetgea ggaggtggge aagaagetgg eegageaett eeceaacaag gecateaagt 2820 tegecaagea cageggegge gaeetggaga teaccaeeca cagetteaac tgeegeggeg 2880 agttetteta etgeaacace egategatee eetgeegeat caageagate atcaacatgt 2940

|            |            |            |            |            | acctgccgca |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | gacaccgaga |      |
|            |            |            |            |            | tacaagtaca |      |
|            |            |            |            |            | accgtgcagg |      |
|            |            |            |            |            | gccatcgagg |      |
|            |            |            |            |            | acccgcgtgc |      |
|            |            |            |            |            | aacatgacct |      |
| ggatgcagtg | ggaccgcgag | atcagcaact | acaccaacac | catctaccgc | ctgctggagg | 3420 |
|            |            |            |            |            | aagtggcaga |      |
|            |            |            |            |            | atccagatct |      |
|            |            |            |            |            | ttccttgacc |      |
| ctggaaggtg | ccactcccac | tgtcctttcc | taataaaatg | aggaaattgc | atcgcattgt | 3660 |
|            |            |            |            |            | gggggaggat |      |
| tgggaagaca | atagcaggca | tgctggggat | gcggtgggct | ctatgggtac | ccaggtgctg | 3780 |
| aagaattgac | ccggttcctc | ctgggccaga | aagaagcagg | cacatcccct | tctctgtgac | 3840 |
|            |            |            |            |            | ctcatagete |      |
| aggagggctc | cgccttcaat | cccacccgct | aaagtacttg | gagcggtctc | tccctccctc | 3960 |
| atcagcccac | caaaccaaac | ctagcctcca | agagtgggaa | gaaattaaag | caagataggc | 4020 |
|            |            |            |            |            | agagaaatca |      |
|            |            |            |            |            | tcctcgctca |      |
|            |            |            |            |            | tcaaaggcgg |      |
|            |            |            |            |            | gcaaaaggcc |      |
|            |            |            |            |            | aggeteegee |      |
|            |            |            |            |            | ccgacaggac |      |
|            |            |            |            |            | gttccgaccc |      |
|            |            |            |            |            | ctttctcata |      |
|            |            |            |            |            | ggctgtgtgc |      |
|            |            |            |            |            | cttgagtcca |      |
|            |            |            |            |            | attagcagag |      |
|            |            |            |            |            | ggctacacta |      |
|            |            |            |            |            | aaaagagttg |      |
|            |            |            |            |            | gtttgcaagc |      |
|            |            |            |            |            | tctacggggt |      |
|            |            |            |            |            | ttatcaaaaa |      |
|            |            |            |            |            | taaagtatat |      |
|            |            |            |            |            | atctcagcga |      |
|            |            |            |            |            | tgaggtctgc |      |
|            |            |            |            |            | tccagccaga |      |
|            |            |            |            |            | gtgattttga |      |
|            |            |            |            |            | tgatccttca |      |
|            |            |            |            |            | gcgtaatgct |      |
|            |            |            |            |            | gcatcaaatg |      |
|            |            |            |            |            | gccgtttctg |      |
|            |            |            |            |            | ggtatcggtc |      |
|            |            |            |            |            | caaaaataag |      |
|            |            |            |            |            | gcaaaagctt |      |
|            |            |            |            |            | caaaatcact |      |
|            |            |            |            |            | atacgcgatc |      |
|            |            |            |            |            | acactgccag |      |
|            |            |            |            |            | atgctgtttt |      |
|            |            |            |            |            | aatgcttgat |      |
|            |            |            |            |            | ctgtaacatc |      |
|            |            |            |            |            | gcttcccata |      |
|            |            |            |            |            | tatacccata |      |
|            |            |            |            |            | cccgttgaat |      |
|            | cccacgccgg | uacceaaccy | cyycologay | caagacgccl | cocyclyaat | 3240 |

```
atggctcata acaccccttg tattactgtt tatgtaagca gacagtttta ttgttcatga 6300
tgatatattt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc 6360
ecceccece cattattgaa geatttatea gggttattgt eteatgageg gatacatatt 6420
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 6480
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 6540
gaggcccttt cgtc
<210> 102
<211> 6671
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 102
tegegegettt eggegatgae ggegaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatq 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgacg taagtacege ctatagacte tataggeaca eccetttgge 1080
tettatgeat getatactgt ttttggettg gggeetatac accecegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae aeggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
```

cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220

| tggagcagat | gcaccaggac | atcatcagcc | tgtgggacca | gagcctgaag | ccctgcgtga | 2280 |
|------------|------------|------------|------------|------------|------------|------|
| agctgacccc | cctgtgcgtg | accctgcact | gcagcaaccg | atcgaagaac | tgcagcttca | 2340 |
| acatgaccac | cgaggtgcgc | gacaagcgcg | agaaggtgca | cgccctgttc | taccgcctgg | 2400 |
| acatcgtgcc | cctgaagaac | gagagcagca | acaccagcgg | cgactaccgc | ctgatcaact | 2460 |
| gcaacaccag | cgccatcacc | caggcctgcc | ccaaggtgag | cttcgacccc | atccccatcc | 2520 |
| actactgcgc | ccccgccggc | tacgccatcc | tgaagtgcaa | caacaagacc | ttcaacggca | 2580 |
| ccggcccctg | caacaacgtg | agcaccatcc | agtgcaccca | cggcaccaag | cccgtggtga | 2640 |
|            |            |            |            | gatcatcatc |            |      |
|            |            |            |            | cgagagcgtg |            |      |
|            |            |            |            | ctggaacaag |            |      |
| aggtgggcaa | gaagctggcc | gagcacttcc | ccaacaaggc | catcaagttc | gccaagcaca | 2880 |
| gcggcggcga | cctggagatc | accacccaca | gcttcaactg | ccgcggcgag | ttcttctact | 2940 |
| gcaacaccag | cagcctgttc | aacagcacct | acacccccaa | cagcaccgag | aacatcaccg | 3000 |
| gcaccgagaa | cagcatcatc | accatcccct | gccgcatcaa | gcagatcatc | aacatgtggc | 3060 |
|            |            |            |            | catcctgacc |            |      |
|            |            |            |            | catgcacgac |            |      |
|            |            |            |            | cgagctgtac |            |      |
| tggtggagat | caagcccctg | ggcatcgccc | ccaccaaggc | caagettace | gtgcaggtgc | 3300 |
| gccagctgct | gagcggcatc | gtgcagcagc | agagcaacct | gctgcgcgcc | atcgaggccc | 3360 |
| agcagcacat | gctgcagctg | accgtgtggg | gcatcaagca | gctgcagacc | cacatactaa | 3420 |
| ccatcgagcg | ctacctgcgc | gaccagcagc | tcgaggacat | ctggaacaac | atgacctgga | 3480 |
| tgcagtggga | ccgcgagatc | agcaactaca | ccaacaccat | ctaccgcctg | ctqqaqqaca | 3540 |
| gccagaacca | gcaggagaag | aacgagcagg | acctgctggc | cctggacaag | tggcagaacc | 3600 |
| tgtggacctg | gttcggcatc | accaactggc | tgtggtgaac | acgtgggatc | cagatetget | 3660 |
| gtgccttcta | gttgccagcc | atctgttgtt | tgcccctccc | ccgtgccttc | cttgaccctg | 3720 |
| gaaggtgcca | ctcccactgt | cctttcctaa | taaaatgagg | aaattgcatc | gcattgtctg | 3780 |
| agtaggtgtc | attctattct | ggggggtggg | gtggggcagg | acagcaaggg | ggaggattgg | 3840 |
| gaagacaata | gcaggcatgc | tggggatgcg | gtgggctcta | tgggtaccca | ggtgctgaag | 3900 |
| aattgacccg | gttcctcctg | ggccagaaag | aagcaggcac | atccccttct | ctgtgacaca | 3960 |
| ccctgtccac | gcccctggtt | cttagttcca | gccccactca | taggacactc | ataġctcagg | 4020 |
|            |            |            |            | cggtctctcc |            |      |
| agcccaccaa | accaaaccta | gcctccaaga | gtgggaagaa | attaaagcaa | gataggctat | 4140 |
| taagtgcaga | gggagagaaa | atgcctccaa | catgtgagga | agtaatgaga | gaaatcatag | 4200 |
| aattttaagg | ccatgattta | aggccatcat | ggccttaatc | ttccgcttcc | tcgctcactg | 4260 |
| actcgctgcg | ctcggtcgtt | cggctgcggc | gagcggtatc | agctcactca | aaggcggtaa | 4320 |
|            |            |            |            | catgtgagca |            |      |
| aaaaggccag | gaaccgtaaa | aaggccgcgt | tgctggcgtt | tttccatagg | ctccgcccc  | 4440 |
| ctgacgagca | tcacaaaaat | cgacgctcaa | gtcagaggtg | gcgaaacccg | acaggactat | 4500 |
| aaagatacca | ggcgtttccc | cctggaagct | ccctcgtgcg | ctctcctgtt | ccgaccctgc | 4560 |
| cgcttaccgg | atacctgtcc | gcctttctcc | cttcgggaag | cgtggcgctt | tctcatagct | 4620 |
| cacgctgtag | gtatctcagt | tcggtgtagg | tcgttcgctc | caagctgggc | tgtgtgcacg | 4680 |
| aaccccccgt | tcagcccgac | cgctgcgcct | tatccggtaa | ctatcgtctt | gagtccaacc | 4740 |
| cggtaagaca | cgacttatcg | ccactggcag | cagccactgg | taacaggatt | agcagagcga | 4800 |
| ggtatgtagg | cggtgctaca | gagttcttga | agtggtggcc | taactacggc | tacactagaa | 4860 |
| gaacagtatt | tggtatctgc | gctctgctga | agccagttac | cttcggaaaa | agagttggta | 4920 |
| gctcttgatc | cggcaaacaa | accaccgctg | gtagcggtgg | tttttttgtt | tgcaagcagc | 4980 |
| agattacgcg | cagaaaaaaa | ggatctcaag | aagatccttt | gatcttttct | acggggtctg | 5040 |
| acgctcagtg | gaacgaaaac | tcacgttaag | ggattttggt | catgagatta | tcaaaaagga | 5100 |
| tcttcaccta | gatcctttta | aattaaaaat | gaagttttaa | atcaatctaa | agtatatatg | 5160 |
| agtaaacttg | gtctgacagt | taccaatgct | taatcagtga | ggcacctatc | tcagcgatct | 5220 |
| gtctatttcg | ttcatccata | gttgcctgac | tcgggggggg | ggggcgctga | ggtctgcctc | 5280 |
| gtgaagaagg | tgttgctgac | tcataccagg | cctgaatcgc | cccatcatcc | agccagaaag | 5340 |
| tgagggagcc | acggttgatg | agagctttgt | tgtaggtgga | ccagttggtg | attttgaact | 5400 |
| tttgctttgc | cacggaacgg | tctgcgttgt | cgggaagatg | cgtgatctga | tccttcaact | 5460 |
| cagcaaaagt | tcgatttatt | caacaaagcc | gccgtcccgt | caagtcagcg | taatgctctg | 5520 |
|            |            |            |            |            |            |      |

```
ccagtgttac aaccaattaa ccaattctga ttagaaaaac tcatcgagca tcaaatgaaa 5580
ctgcaattta ttcatatcag gattatcaat accatatttt tgaaaaagcc gtttctgtaa 5640
tgaaggagaa aactcaccga ggcagttcca taggatggca agatcctggt atcgqtctqc 5700
gattccgact cgtccaacat caatacaacc tattaatttc ccctcgtcaa aaataaggtt 5760
atcaagtgag aaatcaccat gagtgacgac tgaatccggt gagaatggca aaagcttatg 5820
cattlettte cagaettgtt caacaggeca gecattaege tegteateaa aateaetege 5880
atcaaccaaa ccgttattca ttcgtgattg cgcctgagcg agacgaaata cgcgatcgct 5940
gttaaaagga caattacaaa caggaatcga atgcaaccgg cgcaggaaca ctgccagcgc 6000
atcaacaata ttttcacctg aatcaggata ttcttctaat acctggaatg ctgttttccc 6060
ggggatcgca gtggtgagta accatgcatc atcaggagta cggataaaat gcttgatggt 6120
cggaagaggc ataaattccg tcagccagtt tagtctgacc atctcatctg taacatcatt 6180
ggcaacgcta cctttgccat gtttcagaaa caactctggc gcatcgggct tcccatacaa 6240
togatagatt gtogcacctg attgcccgac attatcgcga gcccatttat acccatataa 6300
atcagcatec atgitggaat ttaategegg cetegageaa gaegttteee gttgaatatg 6360
getcataaca eccettgtat taetgtttat gtaageagae agttttattg tteatqatqa 6420
tatattttta tottgtgcaa tgtaacatca gagattttga gacacaacgt ggctttcccc 6480
cccccccat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 6540
atgtatttag aaaaataaac aaataggggt teegegeaca ttteeeegaa aaqtqeeace 6600
tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacqaq 6660
```

gccctttcgt c <210> 103 <211> 6602 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgeg gtgtgaaata cegeacagat gegtaaggag aaaatacege atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegthta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cotataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac totgtootto agagactgac acggactotg tattttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 WO 02/32943 PCT/US01/25721 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440

|            |            |            | 1:         | L          |            | 2440 |
|------------|------------|------------|------------|------------|------------|------|
|            |            | gcgtgggatc |            |            |            |      |
|            |            | cggcggagct |            |            |            |      |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620 |
|            |            | ggctcgcacg |            |            |            |      |
|            |            | ctgagttgtt |            |            |            |      |
|            |            | ggagggcagt |            |            |            |      |
|            |            | ctgacagact |            |            |            |      |
|            |            | cgtgtgatca |            |            |            |      |
|            |            | cagcgcaact |            |            |            |      |
| ccatgegege | ggagggcatc | tacaacgtgc | ggaageageg | gtggatotgg | ototactaco | 2040 |
|            |            |            |            |            |            |      |
|            |            | gccaagacca |            |            |            |      |
|            |            | gtgtgggcca |            |            |            |      |
|            |            | aacgtgaccg |            |            |            |      |
|            |            | atcatcagcc |            |            |            |      |
|            |            | accctgcact |            |            |            |      |
|            |            | gacaagcgcg |            |            |            |      |
|            |            | gagagcagca |            |            |            |      |
|            |            | caggcctgcc |            |            |            |      |
| actactgcgc | ccccgccggc | tacgccatcc | tgaagtgcaa | caacaagacc | ttcaacggca | 2580 |
| ccggcccctg | caacaacgtg | agcaccatcc | agtgcaccca | cggcaccaag | cccgtggtga | 2640 |
|            |            | ggcagcctgg |            |            |            |      |
|            |            | accatcatcg |            |            |            |      |
|            |            | tgcaacatca |            |            |            |      |
|            |            | gagcacttcc |            |            |            |      |
|            |            | accacccaca |            |            |            |      |
|            |            | tgccgcatca |            |            |            |      |
|            |            | cccatcgagg |            |            |            |      |
|            |            | ggcggcaccg |            |            |            |      |
|            |            | aactggcgca |            |            |            |      |
|            |            |            |            |            |            |      |
|            |            | cccaccaagg |            |            |            |      |
|            |            | cagagcaacc |            |            |            |      |
|            |            | ggcatcaagc |            |            |            |      |
|            |            | ctcgaggaca |            |            |            |      |
|            |            | accaacacca |            |            |            |      |
|            |            | gacctgctgg |            |            |            |      |
|            |            | ctgtggtgaa |            |            |            |      |
| agttgccagc | catctgttgt | ttgcccctcc | cccgtgcctt | ccttgaccct | ggaaggtgcc | 3660 |
| actcccactg | tcctttccta | ataaaatgag | gaaattgcat | cgcattgtct | gagtaggtgt | 3720 |
| cattctattc | tggggggtgg | ggtggggcag | gacagcaagg | gggaggattg | ggaagacaat | 3780 |
|            |            | ggtgggctct |            |            |            |      |
|            |            | gaagcaggca |            |            |            |      |
| cgcccctggt | tcttagttcc | agccccactc | ataggacact | catagctcag | gagggctccg | 3960 |
| ccttcaatcc | cacccgctaa | agtacttgga | geggtetete | cctccctcat | cagcccacca | 4020 |
| aaccaaacct | agcctccaag | agtgggaaga | aattaaagca | agataggcta | ttaagtgcag | 4080 |
|            |            | acatgtgagg |            |            |            |      |
|            |            | tggccttaat |            |            |            |      |
|            |            | cgagcggtat |            |            |            |      |
|            |            | gcaggaaaga |            |            |            |      |
|            |            | ttgctggcgt |            |            |            |      |
|            |            | agtcagaggt |            |            |            |      |
|            |            | tecetegtge |            |            |            |      |
|            |            | ccttcgggaa |            |            |            |      |
|            |            | gtcgttcgct |            |            |            |      |
|            |            | ttatccggta |            |            |            |      |
| cccagcccga | cegergegee | ccacccggta | accaccycoc | Lyaycccaac | ccygcaayac | 1000 |

```
acgaettate gecaetggea geagecaetg gtaacaggat tageagageg aggtatgtag 4740
geggtgetae agagttettg aagtggtgge etaactaegg etacactaga agaacaqtat 4800
ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt agctcttgat 4860
ccggcaaaca aaccaccgct ggtagcggtg gttttttttgt ttgcaagcag cagattacgc 4920
gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgctcagt 4980
ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct 5040
agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaactt 5100
ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc 5160
gttcatccat agttgcctga ctcggggggg gggggcgctg aggtctgcct cgtgaagaaq 5220
gtgttgctga ctcataccag gcctgaatcg ccccatcatc cagccagaaa gtgagggagc 5280
cacggttgat gagagetitg ttgtaggtgg accagttggt gattttgaac ttttgctttg 5340
ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg atccttcaac tcagcaaaag 5400
ttogatttat tcaacaaago ogcogtocog tcaagtcago gtaatgotot gccagtgtta 5460
caaccaatta accaattetg attagaaaaa ctcatcgagc atcaaatgaa actgcaattt 5520
attcatatca ggattatcaa taccatattt ttgaaaaagc cgtttctgta atgaaggaga 5580
aaactcaccg aggcagttcc ataggatggc aagatcctgg tatcggtctg cgattccgac 5640
tegtecaaca teaatacaac etattaattt eeeetegtea aaaataaggt tateaaqtqa 5700
gaaatcacca tgagtgacga ctgaatccgg tgagaatggc aaaaqcttat qcatttcttt 5760
ccagacttgt tcaacaggcc agccattacg ctcgtcatca aaatcactcg catcaaccaa 5820
accettatte attegtgatt gegeetgage gagaegaaat acgegatege tettaaaagg 5880
acaattacaa acaggaateg aatgeaaceg gegeaggaae aetgeeageg cateaacaat 5940
attttcacct gaatcaggat attcttctaa tacctggaat gctgttttcc cggggatcgc 6000
agtggtgagt aaccatgcat catcaggagt acggataaaa tgcttgatgg tcggaaqaqq 6060
acctttgcca tgtttcagaa acaactctgg cgcatcgggc ttcccataca atcgatagat 6180
tgtcgcacct gattgcccga cattatcgcg agcccattta tacccatata aatcaqcatc 6240
catgliggaa titaatcgcg gcctcqagca agacgtticc cgttqaatat ggctcataac 6300
acceptigia thacigitha ighaageaga cagithtath githeatgaig atalaithth 6360
atettgtgea atgtaacate agagattttg agacacaacg tggettteee eeceeecea 6420
ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta 6480
gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac ctqacqtcta 6540
agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg 6600
tc
                                                                6602
```

```
<210> 104
<211> 6683
<212> DNA
<213> Artificial Sequence
<223> plasmid pVR1012x/s containing HIV genes
<400> 104
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtq 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattqta ctgagaqtqc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgq 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta cogccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaqt acqccccta ttqacqtcaa 600
```

|   | <b>+</b>   | hamaaaaaat | aaasttabaa | aaaabaaaba | +-+-       |            |      |
|---|------------|------------|------------|------------|------------|------------|------|
|   |            | tggcccgcct |            |            |            |            |      |
|   |            | atctacgtat |            |            |            |            |      |
|   |            | cgtggatagc |            |            |            |            |      |
|   | cgtcaatggg | agtttgtttt | ggcaccaaaa | tcaacgggac | tttccaaaat | gtcgtaacaa | 840  |
|   | ctccgcccca | ttgacgcaaa | tgggcggtag | gcgtgtacgg | tgggaggtct | atataagcag | 900  |
|   |            | gtgaaccgtc |            |            |            |            |      |
|   |            | cgggaccgat |            |            |            |            |      |
|   |            | aagagtgacg |            |            |            |            |      |
|   |            | gctatactgt |            |            |            |            |      |
|   |            | tatagettag |            |            |            |            |      |
|   |            | gatactttcc |            |            |            |            |      |
|   |            | atgccaatac |            |            |            |            |      |
|   |            | ccatttatta |            |            |            |            |      |
|   |            | attaaacata |            |            |            |            |      |
|   |            |            |            |            |            |            |      |
|   |            | tctccggtag |            |            |            |            |      |
|   |            | ggtcgctcgg |            |            |            |            |      |
|   |            | ccaccaccac |            |            |            |            |      |
|   |            | gtggagattg |            |            |            |            |      |
|   |            | atgcaggcag |            |            |            |            |      |
|   |            | tgttaacggt |            |            |            |            |      |
|   |            | gacataatag |            |            |            |            |      |
|   | tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatcgcca | 1920 |
|   | ccatgcgcgt | ggagggcatc | cagcgcaact | ggaagcagtg | gtggatctgg | ggcatcctgg | 1980 |
|   | gcttctggat | ggtgatgatc | tacaacgtgc | gcggcaacct | gtgggtgacc | gtgtactacg | 2040 |
|   | gcgtgcccgt | gtggaaggag | gccaagacca | ccctgttctg | cgccagcgac | gccaaggcct | 2100 |
|   | acgacgccga | ggtgcacaac | gtgtgggcca | cccacgcctg | cgtgcccacc | gaccccaacc | 2160 |
| • |            | ggtgctggag |            |            |            |            |      |
|   |            | gcaccaggac |            |            |            |            |      |
|   |            | cctgtgcgtg |            |            |            |            |      |
|   |            | cgaggtgcgc |            |            |            |            |      |
|   |            | cctgaagaac |            |            |            |            |      |
|   |            | cgccatcacc |            |            |            |            |      |
|   |            | cccgccggc  |            |            |            |            |      |
|   |            | caacaacgtg |            |            |            |            |      |
|   |            | gctgctgaac |            |            |            |            |      |
|   |            | caacgtgaag |            |            |            |            |      |
|   |            | caacaacaac |            |            |            |            |      |
|   |            |            |            |            |            | •          |      |
|   |            | cgagatcatc |            |            |            |            |      |
|   |            | gaccetgeag |            |            |            |            |      |
|   |            | cgccaagcac |            |            |            |            |      |
|   |            | gttcttctac |            |            |            |            |      |
|   |            | gcagggcgtg |            |            |            |            |      |
|   | _          | caacatcacc |            |            |            |            |      |
|   |            | cttccaccc  |            |            |            |            |      |
|   |            | ggtggtggag |            |            |            |            |      |
|   |            | gcgccagctg |            |            |            |            |      |
|   |            | ccagcagcac |            |            |            |            |      |
|   |            | ggccatcgag |            |            |            |            |      |
|   |            | gatgcagtgg |            |            |            |            |      |
|   |            | cagccagaac |            |            |            |            |      |
|   | agtggcagaa | cctgtggacc | tggttcggca | tcaccaactg | gctgtggtga | acacgtggga | 3660 |
|   | tccagatctg | ctgtgccttc | tagttgccag | ccatctgttg | tttgcccctc | ccccgtgcct | 3720 |
|   | tccttgaccc | tggaaggtgc | cactcccact | gtcctttcct | aataaaatga | ggaaattgca | 3780 |
|   | tcgcattgtc | tgagtaggtg | tcattctatt | ctggggggtg | gggtggggca | ggacagcaag | 3840 |
|   |            | gggaagacaa |            |            |            |            |      |
|   |            |            |            |            |            |            |      |

```
caggingetqa aqaatigacc cggitteetee tgggecaqaa aqaaqcagge acateceett 3960
ctctgtgaca caccctgtcc acgcccctgg ttcttagttc cagccccact cataggacac 4020
teatagetea ggagggetee geetteaate ceaceegeta aagtaettgg ageggtetet 4080
ccctccctca tcagcccacc aaaccaaacc tagcctccaa gagtgggaag aaattaaagc 4140
aagataqqct attaagtgca gagggagaga aaatgcctcc aacatgtgag gaagtaatga 4200
qaqaaatcat aqaattttaa qqccatqatt taaqqccatc atqqccttaa tcttccqctt 4260
cetegeteac tgactegetg egeteggteg tteggetgeg gegageggta teageteact 4320
caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag 4380
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 4440
ggeteegeee eeetgaegag cateacaaaa ategaegete aagteagagg tggegaaace 4500
egacaggact ataaagatac caggegttte cocetggaag etecetegtg egeteteetg 4560
tteegaceet geegettace ggatacetgt cegeetttet eeetteggga agegtggege 4620
tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg 4680
getgtgtgca egaacccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc 4740
ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga 4800
ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg cctaactacg 4860
gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa 4920
aaagagttgg tagetettga teeggeaaac aaaccacege tggtageggt ggtttttttg 4980
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 5040
ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat 5100
tatcaaaaaq qatcttcacc tagatccttt taaattaaaa atqaaqtttt aaatcaatct 5160
aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta 5220
teteagegat etgtetattt egtteateea tagttgeetg actegggggg ggggggget 5280
gaggtetgcc tegtgaagaa ggtgttgctg acteatacea ggcctgaate gccccateat 5340
ccagccagaa agtgagggag ccacggttga tgagagcttt gttgtaggtg gaccagttgg 5400
tgattttgaa cttttgcttt gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct 5460
gatcetteaa eteageaaaa gttegattta tteaacaaag eegeegteee gteaagteag 5520
cgtaatgete tgccagtgtt acaaccaatt aaccaattet gattagaaaa actcategag 5580
catcaaatga aactgcaatt tattcatatc aggattatca ataccatatt tttgaaaaag 5640
ccgtttctgt aatgaaggag aaaactcacc gaggcagttc cataggatgg caagatcctg 5700
gtateggtet gegatteega etegteeaac atcaatacaa eetattaatt teeeetegte 5760
aaaaataagg ttatcaagtg agaaatcacc atgagtgacg actgaatccg gtgagaatgg 5820
caaaagctta tgcatttctt tccagacttg ttcaacaggc cagccattac gctcgtcatc 5880
aaaatcactc gcatcaacca aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa 5940
tacgcgatcg ctgttaaaag gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa 6000
cactgccagc gcatcaacaa tattttcacc tgaatcagga tattcttcta atacctggaa 6060
tgctgttttc ccggggatcg cagtggtgag taaccatgca tcatcaggag tacggataaa 6120
atgcttgatg gtcggaagag gcataaattc cgtcagccag tttagtctga ccatctcatc 6180
tgtaacatca ttggcaacgc tacctttgcc atgtttcaga aacaactctg gcgcatcggg 6240
cttcccatac aatcgataga ttgtcgcacc tgattgcccg acattatcgc gagcccattt 6300
atacccatat aaatcagcat ccatgttgga atttaatcgc ggcctcgagc aagacgtttc 6360
ccgttgaata tggctcataa caccccttgt attactgttt atgtaagcag acagttttat 6420
tqttcatgat qatatatttt tatcttqtqc aatgtaacat caqaqatttt qaqacacaac 6480
gtggctttcc cccccccc attattgaag catttatcag ggttattgtc tcatgagcgg 6540
atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg 6600
aaaagtgcca cctgacgtct aagaaaccat tattatcatg acattaacct ataaaaatag 6660
gcgtatcacg aggccctttc gtc
                                                                  6683
```

<210> 105 <211> 6605

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 105 tegegegtht eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cqtatqttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaaccgtc agategeetg gagacgecat ccacqetqtt ttqacctcca 960 tagaagacac egggacegat ecageeteeg eggeegggaa eggtgeattg gaacgeqgat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eeeetttgge 1080 tettatgeat getatactgt tittggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccqtqcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtqttccqqa 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgegegt ggagggcate cagegeaact ggaagcagtg gtggatetgg ggeateetgg 1980 gettetggat ggtgatgate tacaacgtge geggeaacet gtgggtgace gtgtactacg 2040 gegtgeeegt gtggaaggag geeaagaeea ceetgttetg egeeagegae geeaaggeet 2100 acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc qaccccaacc 2160 cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220 tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agetgacece cetgtgegtg accetgeact geageaaceg caccategae tacaacaace 2340 gcaccgacaa catgggcggc gagatcaaga actgcagctt ccgatcgatc aactgcaaca 2400 ccagegecat cacccaggee tgccccaagg tgagettega ecccatecec atecactact 2460 gegeeecege eggetaegee ateetgaagt geaacaacaa gaeetteaac ggeaeeggee 2520 cctgcaacaa cgtgagcacc atccagtgca cccacggcac caagcccqtg qtgagcaccc 2580 agetgetget gaacggcage etggcegagg aggagateat cateegcage aagaacetga 2640 ccgacaacgt gaagaccate atcgtgcacc tgaacgagag cgtggagate aactgcaccc 2700 gccgatcggc ccactgcaac atcagccgca ccgcctggaa caagaccctg caggaggtgg 2760 gcaagaagct ggccgagcac ttccccaaca aggccatcaa gttcgccaag cacagcggcg 2820 gcgacctgga gatcaccacc cacagettca actgccgcgg cgagttette tactgcaaca 2880 ccagcagect gttcaacage acctacacec ccaacageac egagaacate accggcaceg 2940 agaacagcat catcaccatc ccctgccgca tcaagcagat catcaacatg tqqcaqggcg 3000 tgggccgcgc catgtacgcc cccccatcg agggcatcct gacctgccgc agcaacatca 3060

ceggeetget getgaecege gaeggeggea ceggeatgea egacaecgag atetteegee 3120 ccgagggcgg cgacatgcgc gacaactggc gcagcgagct gtacaagtac aaggtggtgg 3180 agateaagee eetgggeate geecceacea aggeeaaget tacegtgeag gtgegeeage 3240 tgctgagcgg catcgtgcag cagcagagca acctgctgcg cgccatcgag gcccagcagc 3300 acatgctgca gctgaccgtg tggggcatca agcagctgca gacccgcgtg ctggccatcg 3360 agegetacet gegegaceag eagetegagg acatetggaa caacatgace tggatgeagt 3420 gggaccgcga gatcagcaac tacaccaaca ccatctaccg cctgctggag gacagccaga 3480 accagcagga gaagaacgag caggacctgc tggccctgga caagtggcag aacctgtgga 3540 cetggtlegg calcaccaac tggctgtggt gaacacgtgg gatccagate tgctgtgcct 3600 tetagttgcc agccatetgt tgtttgcccc tcccccgtgc cttccttgac cctggaaggt 3660 gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcattg tctgagtagg 3720 tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac 3780 aatagcaggc atgctgggga tgcggtgggc tctatgggta cccaggtgct gaagaattga 3840 cccggttcct cctgggccag aaagaagcag gcacatcccc ttctctgtga cacaccctgt 3900 ccaeqceect qqttettaqt tecaqceeca etcataggae acteataget caggaggget 3960 cegectteaa teecaceege taaagtaett ggageggtet eteecteeet cateageeca 4020 ccaaaccaaa cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagtg 4080 cagagggaga gaaaatgcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt 4140 aaggecatga tttaaggeca teatggeett aatetteege tteetegete aetgaetege 4200 tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt 4260 tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg 4320 ccaggaaccq taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg 4380 agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 4440 accapgegtt teceeetgga ageteeeteg tgegetetee tgtteegace etgeegetta 4500 ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 4560 gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc 4620 cegtteagec egacegetge geettateeg gtaactateg tettgagtec aacceggtaa 4680 gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 4740 taggeggtgc tacagagttc ttgaagtggt ggcctaacta eggctacact agaagaacag 4800 tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 4860 qatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 4920 cgcgcaqaaa aaaaggatct caaqaagatc ctttgatctt ttctacgggg tctgacgctc 4980 agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 5040 cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 5100 cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 5160 ttcgttcatc catagttgcc tgactcgggg ggggggggcg ctgaggtctg cctcgtgaag 5220 aaggtgttgc tgactcatac caggcctgaa tcgccccatc atccagccag aaagtgaggg 5280 agccacggtt gatgagagct ttgttgtagg tggaccagtt ggtgattttg aacttttgct 5340 ttqccacqqa acqqtctqcg ttqtcqqqaa qatqcqtgat ctgatccttc aactcagcaa 5400 aagttegatt tatteaacaa ageegeegte eegteaagte agegtaatge tetgeeagtg 5460 ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa 5520 tttattcata tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg 5580 agaaaactca ccgaggcagt tccataggat ggcaagatcc tggtatcggt ctgcgattcc 5640 gactcgtcca acatcaatac aacctattaa tttcccctcg tcaaaaataa ggttatcaag 5700 tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc 5760 tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 5820 caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 5880 aggacaatta caaacaggaa tegaatgeaa ceggegeagg aacactgeca gegeateaac 5940 aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 6000 cgcagtggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6060 aggeataaat teegteagee agtttagtet gaccatetea tetgtaacat cattggeaac 6120 gctacctttg ccatqtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6180 qattqtcqca cctqattqcc cqacattatc gcgagcccat ttatacccat ataaatcagc 6240 atceatgttg gaatttaate geggeetega geaagaegtt teeegttgaa tatggeteat 6300 aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt 6360

```
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6420 ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6480 ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6540 ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6600 tcgtc 6605
```

<210> 106
<211> 6605
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HTV genes

<400> 106

tegegegett eggegatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 caqcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagaqtqc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300, tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 qqqqtcatta qttcataqcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 cataqtaacq ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttq qcaqtacatc aaqtqtatca tatgccaagt acgccccta ttgacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttqqcaqtac atctacqtat taqtcatcqc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tettatgeat getatactgt ttttggettg gggeetatae acccegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattqqctat atqccaatac tctqtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 eqeaqthttt attaaacata gegtgggate tecaegegaa tetegggtae gtgtteegga 1440 catgggetet beteeggtag eggeggaget becaeateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtgqagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 qcaqaaqaaq atqcaqgcaq ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcaqtcacc qtcqtcqaca cgtgtgatca gatatcgcgg ccqctctaqa gatatcgcca 1920 ccatqcqcqt qqaqqcatc cagcqcaact ggaagcagtg gtggatctgg ggcatcctgg 1980 gettetggat ggtgatgate tacaacgtge geggeaacet gtgggtgace gtgtactaeg 2040 gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100 acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160 cccaqqaqat qqtqctqqaq aacqtgaccq agaacttcaa catqtqgqaq aacqacatqq 2220 tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280

| agctgacccc              | cctgtgcgtg               | accctgcact  | gcagcaaccg   | caccatcgac | tacaacaacc  | 2340 |
|-------------------------|--------------------------|-------------|--------------|------------|-------------|------|
|                         | catgggcggc               |             |              |            |             |      |
| ccagcgccat              | cacccaggcc               | tgccccaagg  | tgagcttcga   | ccccatcccc | atccactact  | 2460 |
|                         | cggctacgcc               |             |              |            |             |      |
|                         | cgtgagcacc               |             |              |            |             |      |
|                         | gaacggcagc               |             |              |            |             |      |
|                         | gaagaccatc               |             |              |            |             |      |
|                         | ccactgcaac               |             |              |            |             |      |
| gcaagaagct              | ggccgagcac               | ttccccaaca  | aggccatcaa   | gttcgccaag | cacaqcqqcq  | 2820 |
| gcgacctgga              | gatcaccacc               | cacagettea  | actgccgcgg   | cgagttcttc | tactgcaaca  | 2880 |
| ccagcagcct              | gttcaacagc               | acctacaccc  | ccaacagcac   | cgagaacatc | accoccacco  | 2940 |
|                         | catcaccatc               |             |              |            |             |      |
| tagaccacac              | catgtacgcc               | cccccatcq   | agggcatcct   | gacctgccgc | agcaacatca  | 3060 |
| ccggcctgct              | gctgacccgc               | qacqqcqqca  | ccqqcatqca   | cgacaccgag | atcttcccc   | 3120 |
|                         | cgacatgcgc               |             |              |            |             |      |
| agatcaagcc              | cctgggcatc               | gccccacca   | aggccaagct   | taccatacaa | atacaccaac  | 3240 |
| tactaaacaa              | catcgtgcag               | cagcagagca  | acctactaca   | caccatcaaa | acceaceac   | 3300 |
| acatoctoca              | gctgaccgtg               | tagagata    | agcagctgca   | gacccgcgtg | ctaaccatca  | 3360 |
|                         | gcgcgaccag               |             |              |            |             |      |
| agaaccacaa              | gatcagcaac               | tacaccaaca  | ccatctacca   | cctactagae | cadacaccada | 3480 |
| accadcadda              | gaagaacgag               | caggacctgc  | taaccetaaa   | caagtggcag | aacctotooa  | 3540 |
| cctaattcaa              | catcaccaac               | taactataat  | gaacacgtgg   | catccagate | tactatacat  | 3500 |
| tctagttgc               | agccatctgt               | totttoccc   | teccecatae   | cttccttcac | catageageat | 3660 |
| greactreea              | ctgtcctttc               | ctaataaaat  | gaggaattg    | catcocatto | tetggaagge  | 3720 |
| totcattcta              | ttctgggggg               | tagagtaga   | caggagagaccg | addddadda  | ttagassasa  | 3720 |
| aataggagg               | atgctgggga               | tacaataaa   | tetatogota   | agggggagga | ccgggaagac  | 3040 |
|                         |                          |             |              |            |             |      |
|                         | cctgggccag<br>ggttcttagt |             |              |            |             |      |
|                         |                          |             |              |            |             |      |
|                         | teccaeeege               |             |              |            |             |      |
|                         | cctagcctcc               |             |              |            |             |      |
|                         | gaaaatgcct               |             |              |            |             |      |
|                         | tttaaggcca               |             |              |            |             |      |
|                         | cgttcggctg               |             |              |            |             |      |
|                         | atcaggggat               |             |              |            |             |      |
|                         | taaaaaggcc               |             |              |            |             |      |
|                         | aaatcgacgc               |             |              |            |             |      |
|                         | tccccctgga               |             |              |            |             |      |
|                         | gtccgccttt               |             |              |            |             |      |
|                         | cagttcggtg               |             |              |            |             |      |
| gagagagath              | cgaccgctgc               | geeccateeg  | graduateg    | cettgagtee | aacccggtaa  | 4080 |
| bagagatas               | atcgccactg               | bbassabaat  | ctggtaacag   | gactageaga | gcgaggtatg  | 4/40 |
|                         | tacagagttc               |             |              |            |             |      |
|                         | ctgcgctctg               |             |              |            |             |      |
|                         | acaaaccacc               |             |              |            |             |      |
|                         | aaaaggatct               |             |              |            |             |      |
|                         | aaactcacgt               |             |              |            |             |      |
|                         | tttaaattaa               |             |              |            |             |      |
|                         | cagttaccaa               |             |              |            |             |      |
|                         | catagttgcc               |             |              |            |             |      |
|                         | tgactcatac               |             |              |            |             |      |
| agecaegget<br>bracesees | gatgagagct               | Ligitigtagg | rggaccagtt   | ggtgattttg | aacttttgct  | 5340 |
| Ligidadgga              | acggtctgcg               | ccgccgggaa  | gatgcgtgat   | ctgatccttc | aactcagcaa  | 5400 |
| aagttcgatt              | tattcaacaa               | agccgccgtc  | ccgtcaagtc   | agcgtaatgc | tctgccagtg  | 5460 |
| LEACAACCAA              | ttaaccaatt               | ccgattagaa  | aaactcatcg   | agcatcaaat | gaaactgcaa  | 5520 |
| tttattcata              | tcaggattat               | caataccata  | cctttgaaaa   | agccgtttct | gtaatgaagg  | 5580 |

```
agaaaactca ccgaggcagt tccataggat ggcaagatcc tggtatcggt ctgcgattcc 5640
gactegteca acateaatac aacetattaa tttcccctcg tcaaaaataa ggttatcaag 5700
tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc 5760
tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 5820
caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 5880
aggacaatta caaacaggaa togaatgcaa coqqoqoaqq aacactqcca qoqoatcaac 5940
aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 6000
cgcagtggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6060
aggeataaat teegteagee agtttagtet gaccatetea tetgtaacat cattggcaac 6120
getacetttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6180
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6240
atccatgttg gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6300
aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt 6360
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6420
ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6480
ttaqaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6540
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6600
tegte
<210> 107
<211> 6536
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgaqaqtqc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgq 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tggqaqqtct atataaqcaq 900
agetegttta gtgaaccgtc agategeetg gagacgecat ceacgetgtt ttgaceteca 960
tagaagacac egggacegat ecagecteeg eggeegggaa eggtqcattq qaacqeqqat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tettatgeat getatactgt tittggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
```

cgcagttttt attaaacata gcgtgggate tecacgcgaa tetegggtac gtgtteegga 1440 catgggetet teteoggtag cggcggaget tecacatecg agecetggte ceatgeetee 1500

| agcggctcat | ggtcgctcgg  | cagctccttg | ctcctaacag | tggaggccag   | acttaggcac  | 1560 |
|------------|-------------|------------|------------|--------------|-------------|------|
| agcacaatgc | ccaccaccac  | cagtgtgccg | cacaaggccg | tagcagtaga   | gtatgtgtct  | 1620 |
| gaaaatgagc | gtggagattg  | ggctcgcacq | gctgacgcag | atggaagact   | taaggcagcg  | 1680 |
| gcagaagaag | atgcaggcag  | ctgagttgtt | gtattctgat | aagagtcaga   | ggtaactccc  | 1740 |
| gttgcggtgc | tgttaacggt  | ggagggcagt | gtagtctgag | cagtactcgt   | tactaccaca  | 1800 |
| cgcgccacca | gacataatag  | ctgacagact | aacagactgt | tectttecat   | ggatcttttc  | 1860 |
| tgcagtcacc | qtcqtcqaca  | catataatca | gatatcgcgg | ccactctaga   | gatatogoda  | 1920 |
| ccatgcgcgt | ggagggcatc  | cagcgcaact | ggaagcagtg | ataaatetaa   | gacatectee  | 1980 |
| gcttctggat | ggtgatgatc  | tacaacotoc | gcggcaacct | ataaataacc   | gtatectegg  | 2040 |
| acatacccat | gtggaaggag  | gccaagacca | ccctgttctg | caccaacasc   | accasaacce  | 2100 |
| acqacqccqa | ggtgcacaac  | gtatagacca | cccacgcctg | cataccasc    | gecaaggeee  | 2160 |
| cccaggagat | ggtgctggag  | aacabaacca | agaacttcaa | catgtgggag   | aaccaccaacc | 2220 |
| tqqaqcaqat | gcaccaggac  | atcatcagc  | tgtgggacca | dadcctdaad   | ccctacatas  | 2220 |
| agctgacccc | cctatacata  | accetocact | gcagcaaccg | caccatcaac   | tacaacaaca  | 2200 |
| gcaccgacaa | catggggggg  | gagatcaaga | actgcagctt | cccatccatc   | acaacaacc   | 2340 |
| ccagcgccat | cacccacacc  | tacccaaga  | tgagcttcga | cccatcac     | adcigcaaca  | 2400 |
| acaccccac  | caactacacc  | atcctgaagt | gcaacaacaa | gaggttgaag   | acceaecaec  | 2460 |
| cctocaacaa | catalaceace | atccagtage | cccaccacac | gacccccaac   | ggeaceggee  | 2520 |
| agetgetget | daacaacaac  | ctagagaga  | cccacggcac | caageeegeg   | grgageacce  | 2580 |
| ccacaacat  | gaacggcagc  | atcatagaga | aggagatcat | cateegeage   | aagaacctga  | 2640 |
| accast     | gaagaccacc  | ategegeace | tgaacgagag | cgcggagate   | aactgcaccc  | 2700 |
| geegaeegge | ccaccycaac  | thananan   | ccgcctggaa | caagaccccg   | caggaggtgg  | 2760 |
| gcaagaagcc | ggccgagcac  | cccccaaca  | aggccatcaa | gcccgccaag   | cacageggeg  | 2820 |
| gegaeeegga | gattactact  | cacagettea | actgccgcgg | cgagttette   | tactgcaaca  | 2880 |
| ccegategat | ceeeegeege  | accaagcaga | tcatcaacat | gtggcagggc   | gtgggccgcg  | 2940 |
| ccatgtaege | ccccccatc   | gagggcatcc | tgacctgccg | cagcaacatc   | accggcctgc  | 3000 |
| cgccgaeceg | cgacggcggc  | accggcatgc | acgacaccga | gatetteege   | cccgagggcg  | 3060 |
| gegaeatgeg | cgacaactgg  | cgcagcgagc | tgtacaagta | caaggtggtg   | gagatcaagc  | 3120 |
| ccctgggcat | cgccccacc   | aaggccaagc | ttaccgtgca | ggtgcgccag   | ctgctgagcg  | 3180 |
| geategegea | geageagage  | aacctgctge | gcgccatcga | ggcccagcag   | cacatgctgc  | 3240 |
| agetgacege | grggggcarc  | aagcagctgc | agacccgcgt | gctggccatc   | gagcgctacc  | 3300 |
| tgegegacea | gcagctcgag  | gacatetgga | acaacatgac | ctggatgcag   | tgggaccgcg  | 3360 |
| agatcagcaa | ctacaccaac  | accatctacc | gcctgctgga | ggacagccag   | aaccagcagg  | 3420 |
| agaagaacga | gcaggacctg  | ctggccctgg | acaagtggca | gaacctgtgg   | acctggttcg  | 3480 |
| gcatcaccaa | ctggctgtgg  | tgaacacgtg | ggatccagat | ctgctgtgcc   | ttctagttgc  | 3540 |
| cagccatctg | ttgtttgccc  | ctccccgtg  | ccttccttga | ccctggaagg   | tgccactccc  | 3600 |
| actgtccttt | cctaataaaa  | tgaggaaatt | gcatcgcatt | gtctgagtag   | gtgtcattct  | 3660 |
| attctggggg | gtggggtggg  | gcaggacagc | aagggggagg | attgggaaga   | caatagcagg  | 3720 |
| catgctgggg | atgcggtggg  | ctctatgggt | acccaggtgc | tgaagaattg   | acceggttce  | 3780 |
| tcctgggcca | gaaagaagca  | ggcacatccc | cttctctgtg | acacaccctg   | tccacgcccc  | 3840 |
| tggttcttag | ttccagcccc  | actcatagga | cactcatagc | tcaggagggc   | tccgccttca  | 3900 |
| atcccacccg | ctaaagtact  | tggagcggtc | tctccctccc | tcatcagccc   | accaaaccaa  | 3960 |
| acctagcctc | caagagtggg  | aagaaattaa | agcaagatag | gctattaagt   | gcagagggag  | 4020 |
| agaaaatgcc | tccaacatgt  | gaggaagtaa | tgagagaaat | catagaattt   | taaggccatg  | 4080 |
| atttaaggcc | atcatggcct  | taatcttccg | cttcctcgct | cactgactcg   | ctgcgctcgg  | 4140 |
| tegttegget | gcggcgagcg  | gtatcagctc | actcaaaggc | ggtaatacgg   | ttatccacag  | 4200 |
| aatcagggga | taacgcagga  | aagaacatgt | gagcaaaagg | ccagcaaaag   | gccaggaacc  | 4260 |
| gtaaaaaggc | cgcgttgctg  | gcgtttttcc | ataggctccg | ccccctgac    | gagcatcaca  | 4320 |
| aaaatcgacg | ctcaagtcag  | aggtggcgaa | acccgacagg | actataaaga   | taccaggcgt  | 4380 |
| ttccccctgg | aagctccctc  | gtgcgctctc | ctgttccgac | cctgccgctt   | accggatacc  | 4440 |
| tgtccgcctt | tctcccttcg  | ggaagcgtgg | cgctttctca | tageteaege   | tgtaggtatc  | 4500 |
| tcagttcggt | gtaggtcgtt  | cgctccaagc | tgggctgtgt | gcacgaaccc   | cccgttcagc  | 4560 |
| ccgaccgctg | cgccttatcc  | ggtaactatc | gtcttgagtc | caacccggta   | agacacgact  | 4620 |
| tatcgccact | ggcagcagcc  | actggtaaca | ggattagcag | agcgaggtat   | gtaggcggtg  | 4680 |
| ctacagagtt | cttgaagtgg  | tggcctaact | acggctacac | tagaagaaca   | gtatttggta  | 4740 |
| tctgcgctct | gctgaagcca  | gttaccttcg | gaaaaagagt | tggtagctct   | tgatccggca  | 4800 |
|            |             | _          |            | - <b>-</b> - |             |      |

```
aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa 4860
aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg 4920
aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc 4980
ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg 5040
acagttacca atgettaate agtgaggeae ctateteage gatetgteta tttegtteat 5100
ccatagitge cigacicggg ggggggggc getgaggiet geetegtgaa gaaggigtig 5160
ctgactcata ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt 5220
tgatgagage tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacgg 5280
aacggtctgc gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat 5340
ttattcaaca aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca 5400
attaaccaat tetgattaga aaaacteate gageateaaa tgaaaetgea atttatteat 5460
atcaggatta tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc 5520
accgaggcag ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc 5580
aacatcaata caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc 5640
accatgagtg acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac 5700
ttgttcaaca ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt 5760
attcattcgt gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt 5820
acaaacagga atcgaatgca accggcgcag gaacactgcc agcgcatcaa caatattttc 5880
acctgaatca ggatattett etaataeetg gaatgetgtt tteeegggga tegeagtggt 5940
gagtaaccat gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa 6000
ttccgtcagc cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt 6060
gccatgtttc agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc 6120
acctgattgc ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt 6180
ggaatttaat cgcggcctcg agcaagacgt ttcccgttga atatggctca taacacccct 6240
tgtattactg tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg 6300
tgcaatgtaa catcagagat tttgagacac aacgtggctt tcccccccc cccattattg 6360-
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 6420
taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac 6480
cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc
```

<211> 6617 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 108 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840

<210> 108

| cteceeeee   | ttanaaan    | tagaaaataa |             |                  |            |      |
|-------------|-------------|------------|-------------|------------------|------------|------|
| agglegetta  | ctgacgcaaa  | tgggcggtag | gegegeaegg  | egggaggtet       | acacaagcag | 900  |
| tagecegeeta | grgaaccgcc  | agategeetg | gagacgecat  | ccaegetgtt       | regaceteca | 960  |
| tagaagacac  | egggacegac  | ccagcctccg | cggccgggaa  | cggtgcattg       | gaacgcggat | 1020 |
| Leccegugee  | aagagtgacg  | taagtaccgc | ctatagactc  | tataggcaca       | cccctttggc | 1080 |
| tettatgeat  | gctatactgt  | ttttggcttg | gggcctatac  | acccccgctt       | ccttatgcta | 1140 |
| taggtgatgg  | tatagettag  | cctataggtg | tgggttattg  | accattattg       | accactcccc | 1200 |
| tattggtgac  | gatactttcc  | attactaatc | cataacatgg  | ctctttgcca       | caactatctc | 1260 |
| tattggctat  | ațgccaatac  | tctgtccttc | agagactgac  | acggactctg       | tatttttaca | 1320 |
| ggatggggtc  | ccatttatta  | tttacaaatt | cacatataca  | acaacgccgt       | cccccgtgcc | 1380 |
| cgcagttttt  | attaaacata  | gcgtgggatc | tccacgcgaa  | tctcgggtac       | gtgttccgga | 1440 |
| catgggctct  | tctccggtag  | cggcggagct | tccacatccg  | agccctggtc       | ccatgcctcc | 1500 |
| agcggctcat  | ggtcgctcgg  | cagctccttg | ctcctaacag  | tggaggccag       | acttaggcac | 1560 |
| agcacaatgc  | ccaccaccac  | cagtgtgccg | cacaaggccg  | tggcggtagg       | gtatgtgtct | 1620 |
| gaaaatgagc  | gtggagattg  | ggctcgcacg | 'gctgacgcag | atggaagact       | taaqqcaqcq | 1680 |
| gcagaagaag  | atgcaggcag  | ctgagttgtt | gtattctgat  | aagagtcaga       | ggtaactccc | 1740 |
| gttgcggtgc  | tgttaacggt  | ggagggcagt | gtagtctgag  | cagtactcgt       | tactaccaca | 1800 |
| cgcqccacca  | gacataatag  | ctgacagact | aacagactgt  | tectttecat       | agatettte  | 1860 |
| tgcagtcacc  | atcatcaaca  | cgtgtgatca | gatatcgcgg  | ccactctaga       | gatatogcoa | 1920 |
| ccatacacat  | ggagggcatc  | cagcgcaact | ggaagcagtg  | atagatetaa       | gacatectaa | 1980 |
| gettetggat  | ggtgatgatc  | tacaacgtgc | gcggcaacct  | atagataacc       | gtgtactacg | 2040 |
| gcatacccat  | atagaagag   | gccaagacca | ccctattcta  | caccaacaac       | accasacact | 2100 |
| acgacgccga  | gatacacaac  | gtgtgggcca | cccacacata  | catacceace       | gecaaggeee | 2160 |
| cccaggagat  | gatactagaa  | aacgtgaccg | agaacttcaa  | catataggaa       | gaccccaacc | 2220 |
| todaggagat  | agageeggag  | atcatcagcc | tatagazaa   | gaggtgag         | aacgacatgg | 2220 |
| agetgaecec  | cetatacata  | accetgeact | cgcgggacca  | gageetgaag       | tomogram   | 2280 |
| agecgaecce  | catagagagag | accetgeace | gcagcaaccg  | caccaccgae       | Lacaacaacc | 2340 |
| gcaccgacaa  | cacgggcggc  | gagatcaaga | tangetteen  | cegategate       | aactgcaaca | 2400 |
| gagagagaga  | cacccaggee  | tgccccaagg | cgagettega  | cccatcccc        | acccactact | 2460 |
| gegeeeeege  | eggetaegee  | atcctgaagt | gcaacaacaa  | gaccttcaac       | ggcaccggcc | 2520 |
| cetgeaacaa  | egugageaee  | atccagtgca | cccacggcac  | caagecegtg       | gtgagcaccc | 2580 |
| agetgetget  | gaacggcagc  | ctggccgagg | aggagatcat  | cateegeage       | aagaacctga | 2640 |
| ccgacaacgt  | gaagaccatc  | atcgtgcacc | tgaacgagag  | cgtggagatc       | aactgcaccc | 2700 |
| geeccaacaa  | caacacccgc  | aagagcatcc | gcatcggccc  | cggccaggcc       | ttctacgcca | 2760 |
| ccggcgagat  | catcggcgac  | atccgccagg | cccactgcaa  | catcagccgc       | accgcctgga | 2820 |
| acaagacccc  | gcaggaggtg  | ggcaagaagc | tggccgagca  | cttccccaac       | aaggccatca | 2880 |
| agttcgccaa  | gcacagcggc  | ggcgacctgg | agatcaccac  | ccacagette       | aactgccgcg | 2940 |
| gcgagttctt  | ctactgcaac  | acccgatcga | tcccctgccg  | catcaagcag       | atcatcaaca | 3000 |
|             |             | gccatgtacg |             |                  |            |      |
|             |             | ctgctgaccc |             |                  |            |      |
| agatetteeg  | ccccgagggc  | ggcgacatgc | gcgacaactg  | gcgcagcgag       | ctgtacaagt | 3180 |
| acaaggtggt  | ggagatcaag  | cccctgggca | tcgccccac   | caaggccaag       | cttaccgtgc | 3240 |
| aggtgcgcca  | gctgctgagc  | ggcatcgtgc | agcagcagag  | caacctgctg       | cgcgccatcg | 3300 |
| aggcccagca  | gcacatgctg  | cagctgaccg | tgtggggcat  | caagcagctg       | cagacccgcg | 3360 |
| tgctggccat  | cgagcgctac  | ctgcgcgacc | agcagctcga  | ggacatctgg.      | aacaacatga | 3420 |
| cctggatgca  | gtgggaccgc  | gagatcagca | actacaccaa  | caccatctac       | cgcctgctgg | 3480 |
| aggacagcca  | gaaccagcag  | gagaagaacg | agcaggacct  | gctggccctg       | gacaagtggc | 3540 |
| agaacctgtg  | gacctggttc  | ggcatcacca | actggctgtg  | gtgaacacgt       | gggatccaga | 3600 |
| tctgctgtgc  | cttctagttg  | ccagccatct | gttgtttgcc  | cctcccccgt       | gccttccttg | 3660 |
| accctggaag  | gtgccactcc  | cactgtcctt | tcctaataaa  | atgaggaaat       | tgcatcgcat | 3720 |
| tgtctgagta  | ggtgtcattc  | tattctgggg | ggtggggtgg  | ggcaggacag       | caagggggaq | 3780 |
| gattgggaag  | acaatagcag  | gcatgctggg | gatgcggtgq  | gctctatggg       | tacccaggtg | 3840 |
| ctgaagaatt  | gacccggttc  | ctcctgggcc | agaaagaagc  | aggcacatcc       | ccttctctat | 3900 |
| gacacaccct  | gtccacgccc  | ctggttctta | gttccagccc  | cactcatagg       | acactcatao | 3960 |
| ctcaggaggg  | ctccgccttc  | aatcccaccc | gctaaagtac  | ttggagcggt       | ctctccctcc | 4020 |
| ctcatcagcc  | caccaaacca  | aacctagcct | ccaagagtgq  | gaagaaatta       | aagcaaqata | 4080 |
| ggctattaag  | tgcagaggga  | gagaaaatgc | ctccaacatq  | tgaggaagta       | atgagagaaa | 4140 |
|             |             | _          | _           | - <del>-</del> - |            |      |

```
tcatagaatt ttaaggccat gatttaaggc catcatggcc ttaatcttcc gcttcctcgc 4200
tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 4260
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 4320
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 4380
gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 4440
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 4500
ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc 4560
atageteaeg etgtaggtat eteagttegg tgtaggtegt tegeteeaag etgggetgtg 4620
tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 4680
ccaacceggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 4740
gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 4800
ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 4860
ttggtagete ttgateegge aaacaaacca eegetggtag eggtggtttt tttgtttgca 4920
agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 4980
ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 5040
aaaggatett cacctagate ettttaaatt aaaaatgaag ttttaaatca atetaaagta 5100
tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag 5160
cgatctgtct atttcgttca tccatagttg cctgactcgg ggggggggg cgctgaggtc 5220
tgcctcgtga agaaggtgtt gctgactcat accaggcctg aatcgcccca tcatccagcc 5280
agaaagtgag ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt 5340
tgaacttttg ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatcct 5400
tcaactcagc aaaagttcga tttattcaac aaagccgccg tcccgtcaag tcagcgtaat 5460
gctctgccag tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa 5520
atgaaactgc aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt 5580
ctgtaatgaa ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg 5640
gtctgcgatt ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat 5700
aaggttatca agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag 5760
cttatgcatt tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc 5820
actogcatca accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg 5880
atcgctgtta aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc 5940
cagcgcatca acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt 6000
tttcccgggg atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaatgctt 6060
gatggtcgga agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac 6120
atcattggca acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc 6180
atacaatcga tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc 6240
atataaatca gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg 6300
aatatggctc ataacacccc, ttgtattact gtttatgtaa gcagacagtt ttattgttca 6360
tgatgatata tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct 6420
trececece ecceattatt gaageattta teagggttat tgteteatga geggatacat 6480
atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 6540
gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat 6600
cacgaggccc tttcgtc
                                                                  6617
```

```
<210> 109
<211> 6713
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 109
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
```

| cadcredec   | gtaagcggat | gccgggagca  | gacaagcccg  | tcagggcgcg    | tcagcgggtg | 120  |
|-------------|------------|-------------|-------------|---------------|------------|------|
| rrggcgggtg  | tcggggctgg | cttaactatg  | cggcatcaga  | gcagattgta    | ctgagagtgc | 180  |
| accatatgcg  | gtgtgaaata | ccgcacagat  | gcgtaaggag  | aaaataccgc    | atcagattgg | 240  |
| ctattggcca  | ttgcatacgt | tgtatccata  | tcataatatg  | tacatttata    | ttggctcatg | 300  |
| tccaacatta  | ccgccatgtt | gacattgatt  | attgactagt  | tattaatagt    | aatcaattac | 360  |
| ggggtcatta  | gttcatagcc | catatatgga  | gttccgcgtt  | acataactta    | cggtaaatgg | 420  |
| cccgcctggc  | tgaccgccca | acgacccccg  | cccattgacg  | tcaataatga    | cgtatgttcc | 480  |
| catagtaacg  | ccaataggga | ctttccattg  | acgtcaatgg  | gtggagtatt    | tacggtaaac | 540  |
| tgcccacttg  | gcagtacatc | aagtgtatca  | tatgccaagt  | acgcccccta    | ttgacgtcaa | 600  |
| tgacggtaaa  | tggcccgcct | ggcattatgc  | ccagtacatq  | accttatggg    | actttcctac | 660  |
| ttggcagtac  | atctacgtat | tagtcatcgc  | tattaccatq  | gtgatgcggt    | tttggcagta | 720  |
| catcaatggg  | cgtggatagc | ggtttgactc  | acqqqqattt  | ccaaqtctcc    | accccattga | 780  |
| cgtcaatggg  | agtttgttt  | ggcaccaaaa  | tcaacqqqac  | tttccaaaat    | atcataacaa | 840  |
| ctccqcccca  | ttgacgcaaa | tagacagtag  | acatatacaa  | taggaagtet    | atataagcag | 900  |
| agctcgttta  | gtgaaccgtc | agategeetg  | gagacgccah  | ccacactatt    | ttgacctcca | 960  |
| tagaagacac  | cgggaccgat | ccadecteca  | Sagacaccac  | coatacatta    | gaacccca   | 1020 |
| tececatace  | aagagtgacg | taagtaccog  | ctatagagta  | tatagggaaga   | gaacycygat | 1020 |
| tettateest  | actatactat | ttttaatta   | gagagatataa | acaggeaea     | cetettigge | 1000 |
| taggtgatg   | gctatactgt | ccccggcccg  | basebbaba   | acceeegett    | cettatgeta | 1140 |
| tattagtag   | tatagettag | attactact   | cyggicatig  | accattattg    | accactecee | 1200 |
| tattggtgat  | gatactttcc | actactaatc  | cataacatgg  | ctctttgcca    | caactatctc | 1260 |
| cattggctat  | atgccaatac | tetgtette   | agagactgac  | acggactetg    | tatttttaca | 1320 |
| ggatggggte  | ccatttatta | tttacaaatt  | cacatataca  | acaacgccgt    | ccccgtgcc  | 1380 |
| cgcagtttt   | attaaacata | gcgtgggatc  | tccacgcgaa  | tctcgggtac    | gtgttccgga | 1440 |
| catgggctct  | tctccggtag | cggcggagct  | tccacatccg  | agccctggtc    | ccatgcctcc | 1500 |
| agcggctcat  | ggtcgctcgg | cagctccttg  | ctcctaacag  | tggaggccag    | acttaggcac | 1560 |
| agcacaatgc  | ccaccaccac | cagtgtgccg  | cacaaggccg  | tggcggtagg    | gtatgtgtct | 1620 |
| gaaaatgagc  | gtggagattg | ggctcgcacg  | gctgacgcag  | atggaagact    | taaggcagcg | 1680 |
| gcagaagaag  | atgcaggcag | ctgagttgtt  | gtattctgat  | aagagtcaga    | ggtaactccc | 1740 |
| gttgcggtgc  | tgttaacggt | ggagggcagt  | gtagtctgag  | cagtactcgt    | tgctgccgcg | 1800 |
| cgcgccacca  | gacataatag | ctgacagact  | aacagactgt  | tcctttccat    | gggtcttttc | 1860 |
| tgcagtcacc  | gtcgtcgaca | cgtgtgatca  | gatatcgcgg  | ccgctctaga    | gatatcgcca | 1920 |
| ccatgcgcgt  | ggagggcatc | cagcgcaact  | ggaagcagtg  | gtggatctgg    | ggcatcctgg | 1980 |
| gcttctggat  | ggtgatgatc | tacaacgtgc  | gcggcaacct  | gtgggtgacc    | gtgtactacg | 2040 |
| gcgtgcccgt  | gtggaaggag | gccaagacca  | ccctgttctg  | cgccagcgac    | gccaaggcct | 2100 |
| acgacgccga  | ggtgcacaac | gtgtgggcca  | cccacgcctg  | cgtgcccacc    | gaccccaacc | 2160 |
| cccaggagat  | ggtgctggag | aacgtgaccg  | agaacttcaa  | catgtgggag    | aacgacatgg | 2220 |
| tggagcagat  | gcaccaggac | atcatcagcc  | tgtgggacca  | gagcctgaag    | ccctqcqtqa | 2280 |
| agctgacccc  | cctgtgcgtg | accetgeact  | gcagcaaccg  | caccatcgac    | tacaacaacc | 2340 |
| gcaccgacaa  | catgggcggc | gagatcaaga  | actgcagctt  | caacatgacc    | accgaggtgc | 2400 |
| gcgacaagcg  | cgagaaggtg | cacgccctgt  | tctaccqcct  | qqacatcqtq    | cccctgaaga | 2460 |
| acgagagcag  | caacaccagc | ggcgactacc  | gcctgatcaa  | ctocaacacc    | agcgccatca | 2520 |
| cccaggcctg  | ccccaaggtg | agettegace  | ccatccccat  | ccactactcc    | acccccacca | 2580 |
| gctacgccat  | cctgaagtgc | aacaacaaga  | ccttcaacgg  | caccaacccc    | tacaacaaca | 2640 |
| tgagcaccat  | ccagtgcacc | cacqqcacca  | agcccatagt  | dadcacccad    | ctactactas | 2700 |
| acggcagcct  | ggccgaggag | gagatcatca  | tccacagaga  | gageacceag    | casacataa  | 2760 |
| agaccatcat  | cgtgcacctg | aacdadadcd  | tagagatcaa  | chacacacac    | castcacaca | 2020 |
| actocaacat  | cagccgcacc | acchagaaca  | agacctaca   | gazagtagag    | agaceggeee | 2020 |
| ccgagcacht  | ccccaacaag | gccatcaact  | tracceance  | 22422552      | gaceteeses | 2000 |
| beaceaces   | cagettease | taccaccaage | anttottoto  | chances       | gacciggaga | 2000 |
| traaragrag  | cagcttcaac | aacaccacca  | ageceteed   | aggarage ==== | ageageetgt | 3000 |
| traccators  | ctacaccccc | aacaycaccg  | ayaacaccac  | eggcaccgag    | aacagcatca | 3060 |
| totaccacco  | ctgccgcatc | gacatactac  | actacaegeg  | gcagggcgcg    | ggccgcgcca | 3120 |
| tracrorror  | ccccatcgag | ggcatcottga | cougeogcag  | caacatcacc    | ggcctgctgc | 3180 |
| acatecescya | cggcggcacc | ageacycaeg  | acaccgagat  | cttccgccc     | yagggcggcg | 3240 |
| tagaanhaan  | caactggcgc | agegagetge  | acaagcacaa  | ggtggtggag    | accaagcccc | 3300 |
| rgggcarcgc  | ccccaccaag | gecaageeta  | cegegeagge  | gcgccagctg    | ccgagcggca | 3360 |

| tcatacaaca | дсададсаас | ctactacaca  | ccatcaaaac | ccaccaccac | atgctgcagc | 2420   |
|------------|------------|-------------|------------|------------|------------|--------|
| tgaccatata | gggcatcaag | cadchdcada  | ccacatast  | ggcgatggac | cgctacctgc | 3400   |
| gcgaccagca | actcaaaaac | atctggaaca  | acatgacetg | ggccaccgag | gaccgcgaga | 3400   |
| tcagcaacta | caccaacacc | atctaccaca  | tactacacce | gacgcagtgg | cagcaggaga | 3540   |
| agaacgagca | ggacctgctg | accetagee   | agtaggagga | cagecagaac | tggttcggca | 3600   |
| teaccaacto | actataataa | acacatagaca | tecagateta | stategatta | tagttgccag | . 3660 |
| ccatctatta | tttacccctc | ccccatacst  | techtarea  | tagaaaataa | cagtigedag | 3720   |
| atcettteet | aataaaataa | ggaaattgg   | tagastata  | tggaaggtge | cactcccact | 3780   |
| ctagagagta | agatagaga  | ggaaaccgca  | aggaagaahh | rgagraggrg | tcattctatt | 3840   |
| actaggata  | caataaasta | tataaataaa  | ggggaggacc | gggaagacaa | tagcaggcat | 3900   |
| taaaccaaa  | 292299000  | actgggtatt  | caggigeiga | agaattgacc | cggttcctcc | 3960   |
| ttettaatta | agaagcaggc | acaccccccc  | testacatas | caccetgtee | acgcccctgg | 4020   |
| ccaccagete | aagtacttgg | accaggacac  | ccatagetea | ggagggetee | gccttcaatc | 4080   |
| tagggtgga  | aagtacttgg | ageggtetet  | coccoccca  | ccagcccacc | aaaccaaacc | 4140   |
| aastacataa | gagugggaag | adattadage  | aagatagget | accaagcgca | gagggagaga | 4200   |
| tangectte  | aacacgcgag | gaagtaatga  | gagaaaccac | agaatttaa  | ggccatgatt | 4260   |
| thoughte   | acggeettaa | tetteegett  | ccccgccac  | rgactcgctg | cgctcggtcg | 4320   |
| cceggeegeg | gegageggta | tcagctcact  | caaaggcggt | aatacggtta | tccacagaat | 4380   |
| caggggataa | cgcaggaaag | aacatgtgag  | caaaaggcca | gcaaaaggcc | aggaaccgta | 4440   |
| aaaaggccgc | gttgctggcg | tttttccata  | ggctccgccc | ccctgacgag | catcacaaaa | 4500   |
| atcgacgctc | aagtcagagg | tggcgaaacc  | cgacaggact | ataaagatac | caggcgtttc | 4560   |
| cccctggaag | ctccctcgtg | cgctctcctg  | ttccgaccct | gccgcttacc | ggatacctgt | 4620   |
| ccgcctttct | cccttcggga | agcgtggcgc  | tttctcatag | ctcacgctgt | aggtatctca | 4680   |
| gttcggtgta | ggtcgttcgc | tccaagctgg  | gctgtgtgca | cgaacccccc | gttcagcccg | 4740   |
| accgctgcgc | cttatccggt | aactatcgtc  | ttgagtccaa | cccggtaaga | cacgacttat | 4800   |
| cgccactggc | agcagccact | ggtaacagga  | ttagcagagc | gaggtatgta | ggcggtgcta | 4860   |
| cagagttctt | gaagtggtgg | cctaactacg  | gctacactag | aagaacagta | tttggtatct | 4920   |
| gcgctctgct | gaagccagtt | accttcggaa  | aaagagttgg | tagctcttga | tccggcaaac | 4980   |
| aaaccaccgc | tggtagcggt | ggttttttg   | tttgcaagca | gcagattacg | cgcagaaaaa | 5040   |
| aaggatctca | agaagatcct | ttgatctttt  | ctacggggtc | tgacgctcag | tggaacgaaa | 5100   |
| actcacgtta | agggattttg | gtcatgagat  | tatcaaaaag | gatcttcacc | tagatccttt | 5160   |
| taaattaaaa | atgaagtttt | aaatcaatct  | aaagtatata | tgagtaaact | tggtctgaca | 5220   |
| gttaccaatg | cttaatcagt | gaggcaccta  | tctcagcgat | ctgtctattt | cgttcatcca | 5280   |
| tagttgcctg | actcgggggg | ggggggcgct  | gaggtctgcc | tcgtgaagaa | ggtgttgctg | 5340   |
| actcatacca | ggcctgaatc | gccccatcat  | ccagccagaa | agtgagggag | ccacggttga | 5400   |
| tgagagcttt | gttgtaggtg | gaccagttgg  | tgattttgaa | cttttgcttt | gccacggaac | 5460   |
| ggtctgcgtt | gtcgggaaga | tgcgtgatct  | gatccttcaa | ctcagcaaaa | gttcgattta | 5520   |
| ttcaacaaag | ccgccgtccc | gtcaagtcag  | cgtaatgctc | tgccagtgtt | acaaccaatt | 5580   |
| aaccaattct | gattagaaaa | actcatcgag  | catcaaatga | aactgcaatt | tattcatatc | 5640   |
| aggattatca | ataccatatt | tttgaaaaag  | ccgtttctgt | aatgaaggag | aaaactcacc | 5700   |
| gaggcagttc | cataggatgg | caagatcctg  | gtatcggtct | gcgattccga | ctcgtccaac | 5760   |
| atcaatacaa | cctattaatt | tcccctcgtc  | aaaaataagg | ttatcaagtg | agaaatcacc | 5820   |
|            |            |             |            |            | tccagacttg |        |
| ttcaacaggc | cagccattac | gctcgtcatc  | aaaatcactc | gcatcaacca | aaccgttatt | 5940   |
| cattcgtgat | tgcgcctgag | cgagacgaaa  | tacgcgatcg | ctgttaaaag | gacaattaca | 6000   |
| aacaggaatc | gaatgcaacc | ggcgcaggaa  | cactgccagc | gcatcaacaa | tattttcacc | 6060   |
| tgaatcagga | tattcttcta | atacctggaa  | tgctgttttc | ccggggatcg | cagtggtgag | 6120   |
| taaccatgca | tcatcaggag | tacggataaa  | atgcttgatg | qtcqqaaqaq | gcataaattc | 6180   |
| cgtcagccag | tttagtctga | ccatctcatc  | tgtaacatca | ttggcaacgc | tacctttgcc | 6240   |
| atgtttcaga | aacaactctg | gcgcatcggq  | cttcccatac | aatcgataga | ttgtcgcacc | 6300   |
| tgattgcccg | acattatege | gagcccattt  | atacccatat | aaatcagcat | ccatgttgga | 6360   |
| atttaatcgc | ggcctcgagc | aagacgtttc  | ccqttqaata | tggctcataa | caccccttgt | 6420   |
| attactgttt | atgtaagcag | acaqttttat  | tottcatoat | gatatatttt | tatcttgtgc | 6480   |
| aatgtaacat | cagagatttt | gagacacaac  | gtagetttee | Secondana  | attattgaag | 6540   |
| catttatcao | gattattata | tcatgagcgg  | atacatarth | gaatgtattt | agaaaaataa | 6600   |
| acaaataggg | attecacaes | catttcccca  | aaaagtacca | cctdacatat | aagaaaccat | 6660   |
|            |            |             | guy        | Julyacytut | uayaaaccac | 0000   |

tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtc

6713

```
<210> 110
<211> 6644
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HTV genes
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teageggeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cqqtaaatqq 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960'
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tettatgeat getatactgt ttttggettg gggeetatac acceegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agetgacccc cetgtgcgtg accetgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catgggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400
gcgacaagcg cgagaaggtg cacgccctgt tctaccgcct ggacatcgtg cccctgaaga 2460
acgagageag caacaccage ggegactace geetgateaa etgeaacace agegeeatea 2520
```

| cccagaccta | ccccaacata | agettegagg | aastaaaast |            | <b></b>    |      |
|------------|------------|------------|------------|------------|------------|------|
| actacacat  | catanastas | ageceegace | ccatccccat | ccactactgc | accccacca  | 2580 |
| taagaagaat | ccegaagege | aacaacaaya | cerceaegg  | caccggcccc | rgcaacaacg | 2640 |
| acaggaagga | agagagaga  | cacggeacea | agcccgtggt | gagcacccag | ctgctgctga | 2700 |
| acggcagcct | ggccgaggag | gagaccatca | Leegeageaa | gaacctgacc | gacaacgtga | 2760 |
| agaccaccac | egtgeacetg | aacgagageg | tggagatcaa | ctgcacccgc | cgatcggccc | 2820 |
| actgcaacat | cageegeace | gcctggaaca | agaccctgca | ggaggtgggc | aagaagctgg | 2880 |
| ccgagcactt | ccccaacaag | gccatcaagt | tcgccaagca | cagcggcggc | gacctggaga | 2940 |
| tcaccaccca | cagcttcaac | tgccgcggcg | agttcttcta | ctgcaacacc | cgatcgatcc | 3000 |
| cctgccgcat | caagcagatc | atcaacatgt | ggcagggcgt | gggccgcgcc | atgtacgccc | 3060 |
| cccccatcga | gggcatcctg | acctgccgca | gcaacatcac | cggcctgctg | ctgacccgcg | 3120 |
| acggcggcac | cggcatgcac | gacaccgaga | tetteegeee | cgagggcggc | gacatgcgcg | 3180 |
| acaactggcg | cagcgagctg | tacaagtaca | aggtggtgga | gatcaagccc | ctgggcatcg | 3240 |
| cccccaccaa | ggccaagctt | accgtgcagg | tgcgccagct | gctgagcggc | atcgtgcagc | 3300 |
| agcagagcaa | cctgctgcgc | gccatcgagg | cccagcagca | catgctgcag | ctgaccgtgt | 3360 |
| ggggcatcaa | gcagctgcag | acccgcgtgc | tggccatcga | gcgctacctg | cgcgaccaqc | 3420 |
| agctcgagga | catctggaac | aacatgacct | ggatgcagtg | ggaccgcgag | atcagcaact | 3480 |
| acaccaacac | catctaccgc | ctgctggagg | acagccagaa | ccagcaggag | aagaacgagc | 3540 |
| aggacctgct | ggccctggac | aagtggcaga | acctgtggac | ctggttcggc | atcaccaact | 3600 |
| ggctgtggtg | aacacgtggg | atccagatct | gctgtgcctt | ctagttgcca | gccatctqtt | 3660 |
| gtttgcccct | ccccgtgcc  | ttccttgacc | ctggaaggtg | ccactcccac | tgtcctttcc | 3720 |
| taataaaatg | aggaaattgc | atcgcattgt | ctgagtaggt | gtcattctat | tctggggggt | 3780 |
| ggggtggggc | aggacagcaa | gggggaggat | tgggaagaca | atagcaggca | tgctggggat | 3840 |
| gcggtgggct | ctatgggtac | ccaggtgctg | aagaattgac | ccggttcctc | ctqqqccaqa | 3900 |
| aagaagcagg | cacatcccct | tctctgtgac | acaccctgtc | cacgcccctg | gttcttagtt | 3960 |
| ccagccccac | tcataggaca | ctcatagctc | aggagggctc | cgccttcaat | cccacccgct | 4020 |
| aaagtacttg | gagcggtctc | tecetecete | atcagcccac | caaaccaaac | ctagcctcca | 4080 |
| agagtgggaa | gaaattaaag | caagataggc | tattaagtgc | agagggagag | aaaatgcctc | 4140 |
| caacatgtga | ggaagtaatg | agagaaatca | tagaatttta | aggccatgat | ttaaggccat | 4200 |
| catggcctta | atcttccgct | tcctcgctca | ctgactcgct | gcgctcggtc | gttcggctgc | 4260 |
| ggcgagcggt | atcagctcac | tcaaaggcgg | taatacggtt | atccacagaa | tcaggggata | 4320 |
| acgcaggaaa | gaacatgtga | gcaaaaggcc | agcaaaaggc | caggaaccgt | aaaaaggccg | 4380 |
| cgttgctggc | gtttttccat | aggctccgcc | cccctgacga | gcatcacaaa | aatcgacgct | 4440 |
| caagtcagag | gtggcgaaac | ccgacaggac | tataaagata | ccaggcgttt | cccctggaa  | 4500 |
| gctccctcgt | gcgctctcct | gttccgaccc | tgccgcttac | cggatacctg | teegeettte | 4560 |
| tcccttcggg | aagcgtggcg | ctttctcata | gctcacgctg | taggtatctc | agttcggtgt | 4620 |
| aggtcgttcg | ctccaagctg | ggctgtgtgc | acgaaccccc | cgttcagccc | gaccgctgcg | 4680 |
| ccttatccgg | taactatcgt | cttgagtcca | acccggtaag | acacgactta | togccactqq | 4740 |
| cagcagccac | tggtaacagg | attagcagag | cgaggtatgt | aggcggtgct | acagagttct | 4800 |
| tgaagtggtg | gcctaactac | ggctacacta | gaagaacagt | atttggtatc | tgcgctctgc | 4860 |
| tgaagccagt | taccttcgga | aaaagagttg | gtagctcttg | atccggcaaa | caaaccaccq | 4920 |
| ctggtagcgg | tggtttttt  | gtttgcaagc | agcagattac | gcgcagaaaa | aaaggatctc | 4980 |
| aagaagatcc | tttgatcttt | tctacggggt | ctgacgctca | gtggaacgaa | aactcacqtt | 5040 |
| aagggatttt | ggtcatgaga | ttatcaaaaa | ggatcttcac | ctagatcctt | ttaaattaaa | 5100 |
| aatgaagttt | taaatcaatc | taaagtatat | atgagtaaac | ttggtctgac | agttaccaat | 5160 |
| gcttaatcag | tgaggcacct | atctcagcga | tctgtctatt | tcgttcatcc | atagttgcct | 5220 |
| gactcggggg | ggggggcgc  | tgaggtctgc | ctcgtgaaga | aggtgttgct | gactcatacc | 5280 |
| aggcctgaat | cgccccatca | tccagccaga | aagtgaggga | gccacggttg | atgagagett | 5340 |
| tgttgtaggt | ggaccagttg | gtgattttga | acttttgctt | tgccacggaa | cggtctqcqt | 5400 |
| tgtcgggaag | atgcgtgatc | tgatccttca | actcagcaaa | agttcgattt | attcaacaaa | 5460 |
| gccgccgtcc | cgtcaagtca | gcgtaatgct | ctgccagtgt | tacaaccaat | taaccaattc | 5520 |
| tgattagaaa | aactcatcga | gcatcaaatg | aaactgcaat | ttattcatat | caggattatc | 5580 |
| aataçcatat | ttttgaaaaa | gccgtttctg | taatgaagga | gaaaactcac | cgaggcagtt | 5640 |
| ccataggatg | gcaagatcct | ggtatcggtc | tgcgattccg | actcgtccaa | catcaataca | 5700 |
| acctattaat | ttcccctcgt | caaaaataag | gttatcaagt | gagaaatcac | catgagtgac | 5760 |
| gactgaatcc | ggtgagaatg | gcaaaagctt | atgcatttct | ttccagactt | gttcaacagg | 5820 |
|            |            |            |            | -          |            |      |

<210> 111
<211> 6725
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

-400× 111

tegegegettt eggegatgae ggegaaaace teegacacat geageteeeg gagaeggeea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacqqtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt titttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaeteece 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

| attacaataa  | tattaaaaat  | ~~~~~~     |            |            | <b>1</b>   |      |
|-------------|-------------|------------|------------|------------|------------|------|
| cacacasacas | gacataatag  | ggagggcagt | gragicigag | cagtactcgt | ractaceaca | 1800 |
| tacaataaca  | gacataatag  | catatastas | aacagactge | tcctttccat | gggtettte  | 1860 |
| ccatacacat  | geegeegaea  | cgcgcgacca | gatategegg | ccgctctaga | gacacegeea | 1920 |
| actteteest  | ggagggcacc  | tagegeaace | ggaagcagtg | gtggatctgg | ggeatectgg | 1980 |
| geeteeggat  | ggcgacgacc  | cacaacgtgc | geggeaacet | gtgggtgacc | gtgtactacg | 2040 |
| acgacacaga  | graggaaggag | gecaagacca | ccctgttctg | cgccagcgac | gccaaggcct | 2100 |
| cocacacaca  | ggtgcacaac  | grgrgggeea | cccacgectg | cgtgcceacc | gaccccaacc | 2160 |
| tecaggagat  | ggtgetggag  | aacgtgaccg | agaacttcaa | catgtgggag | aacgacatgg | 2220 |
| agatasagas  | gcaccaggac  | accatcagee | tgtgggacca | gagcctgaag | ccctgcgtga | 2280 |
| agetgaetee  | cetgtgegtg  | accetgeace | gcagcaaccg | caccatcgac | cacaacaacc | 2340 |
|             |             |            |            | caacatgacc |            |      |
|             |             |            |            | ggacatcgtg |            |      |
| acgagagcag  | caacaccagc  | ggegaetace | gcctgatcaa | ctgcaacacc | agcgccatca | 2520 |
| cccaggeerg  | ccccaaggtg  | agettegace | ccatccccat | ccactactgc | gcccccgccg | 2580 |
| getaegeeat  | cetgaagtge  | aacaacaaga | cctccaacgg | caccggcccc | tgcaacaacg | 2640 |
| cgageaecat  | ccagtgeace  | cacggcacca | agcccgtggt | gagcacccag | ctgctgctga | 2700 |
| acggcagcct  | ggccgaggag  | gagatcatca | tccgcagcaa | gaacctgacc | gacaacgtga | 2760 |
| agaccatcat  | cgtgcacctg  | aacgagagcg | tggagatcaa | ctgcacccgc | cccaacaaca | 2820 |
| acacccgcaa  | gagcatccgc  | atcggccccg | gccaggcctt | ctacgccacc | ggcgagatca | 2880 |
| tcggcgacat  | ccgccaggcc  | cactgcaaca | tcagccgcac | cgcctggaac | aagaccctgc | 2940 |
| aggaggtggg  | caagaagctg  | gccgagcact | tccccaacaa | ggccatcaag | ttcgccáagc | 3000 |
| acagcggcgg  | cgacctggag  | atcaccaccc | acagetteaa | ctgccgcggc | gagttcttct | 3060 |
| actgcaacac  | ccgatcgatc  | ccctgccgca | tcaagcagat | catcaacatg | tggcagggcg | 3120 |
| rgggccgcgc  | catgtacgcc  | cccccatcg  | agggcatcct | gacctgccgc | agcaacatca | 3180 |
| ceggeetget  | gctgacccgc  | gacggcggca | ccggcatgca | cgacaccgag | atcttccgcc | 3240 |
| ccgagggcgg  | cgacatgcgc  | gacaactggc | gcagcgagct | gtacaagtac | aaggtggtgg | 3300 |
|             |             |            |            | taccgtgcag |            |      |
| tgctgagcgg  | catcgtgcag  | cagcagagca | acctgctgcg | cgccatcgag | gcccagcagc | 3420 |
| acatgctgca  | gctgaccgtg  | tggggcatca | agcagctgca | gacccgcgtg | ctggccatcg | 3480 |
| agcgctacct  | gcgcgaccag  | cagctcgagg | acatctggaa | caacatgacc | tggatgcagt | 3540 |
| gggaccgcga  | gatcagcaac  | tacaccaaca | ccatctaccg | cctgctggag | gacagccaga | 3600 |
| accagcagga  | gaagaacgag  | caggacctgc | tggccctgga | caagtggcag | aacctgtgga | 3660 |
| cctggttcgg  | catcaccaac  | tggctgtggt | gaacacgtgg | gatecagate | tgctgtgcct | 3720 |
| tctagttgcc  | agccatctgt  | tgtttgcccc | tcccccgtgc | cttccttgac | cctggaaggt | 3780 |
| gccactccca  | ctgtcctttc  | ctaataaaat | gaggaaattg | catcgcattg | tctgagtagg | 3840 |
| tgtcattcta  | ttctgggggg  | tggggtgggg | caggacagca | agggggagga | ttgggaagac | 3900 |
| aatagcaggc  | atgctgggga  | tgcggtgggc | tctatgggta | cccaggtgct | gaagaattga | 3960 |
| cccggttcct  | cctgggccag  | aaagaagcag | gcacatcccc | ttctctgtga | cacaccctgt | 4020 |
| ccacgcccct  | ggttcttagt  | tccagcccca | ctcataggac | actcatagct | caggagggct | 4080 |
| ccgccttcaa  | tcccacccgc  | taaagtactt | ggagcggtct | ctccctccct | catcagccca | 4140 |
| ccaaaccaaa  | cctagcctcc  | aagagtggga | agaaattaaa | gcaagatagg | ctattaagtg | 4200 |
| cagagggaga  | gaaaatgcct  | ccaacatgtg | aggaagtaat | gagagaaatc | atagaatttt | 4260 |
| aaggccatga  | tttaaggcca  | tcatggcctt | aatcttccgc | ttcctcgctc | actgactcgc | 4320 |
| tgcgctcggt  | cgttcggctg  | cggcgagcgg | tatcagctca | ctcaaaggcg | gtaatacggt | 4380 |
| tatccacaga  | atcaggggat  | aacgcaggaa | agaacatgtg | agcaaaaggc | cagcaaaagg | 4440 |
| ccaggaaccg  | taaaaaggcc  | gcgttgctgg | cgtttttcca | taggctccgc | cccctgacg  | 4500 |
| agcatcacaa  | aaatcgacgc  | tcaagtcaga | ggtggcgaaa | cccgacagga | ctataaaqat | 4560 |
| accaggcgtt  | tccccctgga  | ageteceteg | tgcgctctcc | tgttccgacc | ctgccqctta | 4620 |
| ccggatacct  | gtccgccttt  | ctcccttcgg | gaagegtgge | gctttctcat | agctcacqct | 4680 |
| gtaggtatct  | cagttcggtg  | taggtcgttc | gctccaagct | gggctgtgtg | cacgaacccc | 4740 |
| ccgttcagcc  | cgaccgctgc  | gccttatccq | gtaactatco | tcttgagtcc | aacccqqtaa | 4800 |
| gacacgactt  | atcgccactq  | gcagcagcca | ctggtaacag | gattagcaga | gcgaggtatg | 4860 |
| taggcggtgc  | tacagagttc  | ttgaagtgat | ggcctaacta | cggctacact | agaagaacag | 4920 |
| tatttggtat  | ctgcgctctq  | ctgaagccag | ttaccttcaa | aaaaagagtt | ggtagctctt | 4980 |
| gatccggcaa  | acaaaccacc  | gctggtagcq | gtggtttttt | tgtttgcaag | cagcagatta | 5040 |
|             |             |            |            | J 9 9      |            |      |

```
cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 5100
agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 5160
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 5220
cttggtctga cagttaccaa tgcttaatca qtgaggcacc tatctcagcg atctgtctat 5280
ttegtteate catagttgee tgactegggg ggggggggg ctgaggtetg cetegtgaag 5340
aaggtgttgc tgactcatac caggcctgaa tcgccccatc atccagccag aaagtgaggg 5400
agccacggtt gatgagagct ttgttgtagg tggaccagtt ggtgattttg aacttttgct 5460
ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat ctgatccttc aactcagcaa 5520
aagttegatt tatteaacaa ageegeegte eegtcaagte agegtaatge tetgeeagtg 5580
ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa 5640
tttattcata tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg 5700
agaaaactca ccgaggcagt tccataggat ggcaagatcc tggtatcggt ctgcgattcc 5760
gactogtoca acatcaatac aacctattaa tttcccctcg tcaaaaataa ggttatcaag 5820
tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc 5880
tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 5940
caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 6000
aggacaatta caaacaggaa tegaatgcaa ceggegeagg aacaetgeea gegeateaac 6060
aatattttca cetgaatcag gatattette taatacetgg aatgetgttt teeeggggat 6120
cgcagtggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6180
aggcabaaat teegteagee agtttagtet gaccatetea tetgtaacat cattggcaac 6240
gctacetttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6300
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6360
atccatgttg gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6420
aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt 6480
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6540
ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6600
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6660
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6720
tcgtc
```

```
<211> 6824
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 112
tegegegitt eggigatgae ggigaaaace telgacacat geageteeeg gagaeggica 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
eccgcetgge tgacegeeca acgaceceeg eccattgaeg teaataatga egtatgttee 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcqtaacaa 840
```

<210> 112

| ctccgcccca | ttgacgcaaa | tgggcggtag | gcgtgtacgg | tgggaggtct | atataagcag | 900  |
|------------|------------|------------|------------|------------|------------|------|
| agctcgttta | gtgaaccgtc | agatcgcctg | gagacgccat | ccacgctgtt | ttgacctcca | 960  |
| tagaagacac | cgggaccgat | ccagcctccg | cggccgggaa | cggtgcattg | gaacgcggat | 1020 |
| tccccgtgcc | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca | cccctttggc | 1080 |
| tcttatgcat | gctatactgt | ttttggcttg | gggcctatac | acccccgctt | ccttatgcta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg | tgggttattg | accattattg | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260 |
| tattggctat | atgccaatac | tctgtccttc | agagactgac | acggactctg | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacgccgt | ccccgtgcc  | 1380 |
| cgcagttttt | attaaacata | gcgtgggatc | tccacgcgaa | tctcgggtac | gtgttccgga | 1440 |
| catgggctct | tctccggtag | cggcggagct | tccacatccg | agccctggtc | ccatgcctcc | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact | taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg | 1800 |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | tcctttccat | gggtcttttc | 1860 |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatcgcca | 1920 |
|            |            |            |            |            | ggcatcctgg |      |
|            |            |            |            |            | gtgtactacg |      |
| gcgtgcccgt | gtggaaggag | gccaagacca | ccctgttctg | cgccagcgac | gccaaggcct | 2100 |
| acgacgccga | ggtgcacaac | gtgtgggcca | cccacgcctg | cgtgcccacc | gaccccaacc | 2160 |
|            |            |            |            |            | aacgacatgg |      |
| tggagcagat | gcaccaggac | atcatcagcc | tgtgggacca | gagcctgaag | ccctgcgtga | 2280 |
| agctgacccc | cctgtgcgtg | accctgcact | gcagcaaccg | atcgaagaac | tgcagcttca | 2340 |
|            |            |            |            |            | taccgcctgg |      |
|            |            |            |            |            | ctgatcaact |      |
|            |            |            |            |            | atccccatcc |      |
| actactgcgc | ccccgccggc | tacgccatcc | tgaagtgcaa | caacaagacc | ttcaacggca | 2580 |
| ccggcccctg | caacaacgtg | agcaccatcc | agtgcaccca | cggcaccaag | cccgtggtga | 2640 |
| gcacccagct | gctgctgaac | ggcagcctgg | ccgaggagga | gatcatcatc | cgcagcaaga | 2700 |
| acctgaccga | caacgtgaag | accatcatcg | tgcacctgaa | cgagagcgtg | gagatcaact | 2760 |
| gcacccgccc | caacaacaac | acccgcaaga | gcatccgcat | cggccccggc | caggccttct | 2820 |
| acgccaccgg | cgagatcatc | ggcgacatcc | gccaggccca | ctgcaacatc | agccgcaccg | 2880 |
| cctggaacaa | gaccctgcag | gaggtgggca | agaagctggc | cgagcacttc | cccaacaagg | 2940 |
| ccatcaagtt | cgccaagcac | agcggcggcg | acctggagat | caccacccac | agcttcaact | 3000 |
| gccgcggcga | gttcttctac | tgcaacacca | gcagcctgtt | caacagcacc | tacaccccca | 3060 |
| acagcaccga | gaacatcacc | ggcaccgaga | acagcatcat | caccatcccc | tgccgcatca | 3120 |
| agcagatcat | caacatgtgg | cagggcgtgg | gccgcgccat | gtacgccccc | cccatcgagg | 3180 |
| gcatcctgac | ctgccgcagc | aacatcaccg | gcctgctgct | gacccgcgac | ggcggcaccg | 3240 |
| gcatgcacga | caccgagatc | ttccgccccg | agggcggcga | catgcgcgac | aactggcgca | 3300 |
|            |            |            |            |            | cccaccaagg |      |
| ccaagcttac | cgtgcaggtg | cgccagctgc | tgagcggcat | cgtgcagcag | cagagcaacc | 3420 |
|            |            |            |            |            | ggcatcaagc |      |
| agctgcagac | ccgcgtgctg | gccatcgagc | gctacctgcg | cgaccagcag | ctcgaggaca | 3540 |
| tctggaacaa | catgacctgg | atgcagtggg | accgcgagat | cagcaactac | accaacacca | 3600 |
| tctaccgcct | gctggaggac | agccagaacc | agcaggagaa | gaacgagcag | gacctgctgg | 3660 |
|            |            |            |            |            | ctgtggtaca |      |
| tcaagatctt | catcaagatc | gtgggcggcc | tgatcggcct | gcgcatcatc | ttcgccgtgc | 3780 |
| tgagcatctg | aacacgtggg | atccagatct | gctgtgcctt | ctagttgcca | gccatctgtt | 3840 |
|            |            |            |            |            | tgtcctttcc |      |
|            |            |            |            |            | tctggggggt |      |
|            |            |            |            |            | tgctggggat |      |
| gcggtgggct | ctatgggtac | ccaggtgctg | aagaattgac | ccggttcctc | ctgggccaga | 4080 |
| aagaagcagg | cacatcccct | tctctgtgac | acaccctgtc | cacgcccctg | gttcttagtt | 4140 |
|            |            |            |            |            |            |      |

```
ccageceeae teataggaca eteatagete aggagggete egeetteaat eccaeeeget 4200
agagtgggaa gaaattaaag caagataggc tattaagtgc agagggagag aaaatgcctc 4320
caacatgtga ggaagtaatg agagaaatca tagaatttta aggccatgat ttaaggccat 4380
catggcctta atcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc 4440
ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata 4500
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 4560
cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct 4620
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa 4680
getecetegt gegetetect gtteegacee tgeegettae eggatacetg teegeettte 4740
tecetteggg aagegtggeg ettteteata geteaegetg taggtatete agtteggtgt 4800
agglegiteg ciccaagetg ggetgitge acgaaceee egiteageee gacegetgeg 4860
cettateegg taactategt ettgagteea acceggtaag acacgaetta tegecaetgg 4920
cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 4980
tgaagtggtg gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc 5040
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 5100
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 5160
aagaagatee titgatetit tetaeggggt etgaegetea gtggaaegaa aacteaegtt 5220
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 5340
gettaateag tgaggeacet ateteagega tetgtetatt tegtteatee atagttgeet 5400
gactcggggg gggggggcgc tgaggtctgc ctcgtgaaga aggtgttgct gactcatacc 5460
aggcetgaat egececatea tecagecaga aagtgaggga gecaeggttg atgagagett 5520
tgttgtaggt ggaccagttg gtgattttga acttttgctt tgccacggaa cggtctgcgt 5580
tgtcgggaag atgcgtgatc tgatccttca actcagcaaa agttcgattt attcaacaaa 5640
gccgccgtcc cgtcaagtca gcgtaatgct ctgccagtgt tacaaccaat taaccaattc 5700
tgattagaaa aactcatcga gcatcaaatg aaactgcaat ttattcatat caggattatc 5760
aataccatat ttttgaaaaa gccgtttctg taatgaagga gaaaactcac cgaggcagtt 5820
ccataggatg gcaagatcct ggtatcggtc tgcgattccg actcgtccaa catcaataca 5880
acctattaat ttcccctcgt caaaaataag gttatcaagt gagaaatcac catgagtgac 5940
gactgaatcc ggtgagaatg gcaaaagctt atgcatttct ttccagactt gttcaacagg 6000
ccagccatta cgctcgtcat caaaatcact cgcatcaacc aaaccgttat tcattcgtga 6060
ttgcgcctga gcgagacgaa atacgcgatc gctgttaaaa ggacaattac aaacaggaat 6120
cgaatgcaac cggcgcagga acactgccag cgcatcaaca atattttcac ctgaatcagg 6180
atattettet aatacetgga atgetgtttt eeeggggate geagtggtga gtaaceatge 6240
atcatcagga gtacggataa aatgcttgat ggtcggaaga ggcataaatt ccgtcagcca 6300
gtttagtctg accatctcat ctgtaacatc attggcaacg ctacctttgc catgtttcag 6360
aaacaactet ggegeategg getteecata caategatag attgtegeac etgattgeec 6420
gacattatcg cgagcccatt tatacccata taaatcagca tccatgttgg aatttaatcg 6480
eggeetegag caagaegttt ecegttgaat atggeteata acaeceettg tattactgtt 6540
tatgtaagca gacagtttta ttgttcatga tgatatattt ttatcttgtg caatgtaaca 6600
tcagagattt tgagacacaa cgtggctttc cccccccc cattattgaa gcatttatca 6660
gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 6720
ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat 6780
gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtc
                                                               6824
```

```
<210> 113
<211> 6695
<212> DNA
<213> Artificial Sequence
```

<223> plasmid pVR1012x/s containing HTV genes

<400> 113 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcqtt acataactta cqqtaaatqq 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgacg taagtacege ctatagacte tataggeaca eccetttgge 1080 tottatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140 taggtgatgg tatagettag ectataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg ageeetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt teetttecat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980 gettetggat ggtgatgate tacaacgtge geggeaacet gtgggtgace gtgtactacg 2040 gegtgecegt gtggaaggag gecaagacea ceetgttetg egecagegae gecaaggeet 2100 acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160 cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220 tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agetgacece cetgtgegtg accetgeact geageaaceg ategateaac tgcaacacea 2340 gegecateae ccaggeetge eccaaggtga gettegacee catececate cactactgeg 2400 cccccgccgg ctacgccatc ctgaagtgca acaacaagac cttcaacggc accggcccct 2460 gcaacaacgt gagcaccate cagtgcacce acggcaccaa gcccgtggtg agcacccagc 2520 tgctgctgaa cggcagcctg gccgaggagg agatcatcat ccgcaqcaag aacctgaccg 2580 acaacgtgaa gaccatcatc gtgcacctga acgagagcgt ggagatcaac tgcacccgcc 2640 ccaacaacaa caccegcaag agcateegca teggeeeegg ccaggeette tacgecaceg 2700 gcgagatcat cggcgacatc cgccaggccc actgcaacat cagccgcacc gcctggaaca 2760 agaccetgea ggaggtggge aagaagetgg eegageactt eeceaacaag gecateaagt 2820 tegecaagea cageggegge gaeetggaga teaceaecea cagetteaae tgeegeggeg 2880 agttetteta etgeaacace ageageetgt teaacageac etacaceece aacageaceg 2940 agaacatcac eggcacegag aacagcatca teaccateee etgeegeate aagcagatea 3000 tcaacatgtg gcagggcgtg ggccgcgcca tgtacgccc ccccatcgag ggcatcctga 3060 cetgeegeag caacateace ggcetgetge tgaccegega eggeggeace ggcatgeacg 3120

acaccgagat cttccgcccc gagggcggcg acatgcgcga caactggcgc agcgagctgt 3180

|             |                          |             |             |            |            | 2242 |
|-------------|--------------------------|-------------|-------------|------------|------------|------|
|             | ggtggtggag               |             |             |            |            |      |
|             | gcgccagctg               |             |             |            |            |      |
|             | ccagcagcac               |             |             |            |            |      |
|             | ggccatcgag               |             |             |            |            |      |
| acatgacctg  | gatgcagtgg               | gaccgcgaga  | tcagcaacta  | caccaacacc | atctaccgcc | 3480 |
|             | cagccagaac               |             |             |            |            |      |
|             | cctgtggacc               |             |             |            |            |      |
|             | cgtgggcggc               |             |             |            |            |      |
|             | gatccagatc               |             |             |            |            |      |
|             | cttccttgac               |             |             |            |            |      |
| gaggaaattg  | catcgcattg               | tctgagtagg  | totcattcta  | ttctaaaaaa | tagaatagaa | 3840 |
| 2022acacca  | agggggagga               | ttaggaagac  | aatagcaggc  | atactagga  | tacaataaac | 3900 |
| totatooota  | cccaggtgct               | gaagaattga  | cccaattcct  | cctagaccaa | aaagaaggag | 3960 |
| ggagatgggga | ttctctgtga               | cacaccctgt  | ccagggccct  | aghtchtagh | hocagooca  | 4020 |
| gtatattcccc | actcatagct               | cacacacacac | ccaccettcaa | teceacease | taaagtactt | 4080 |
|             | ctccctccct               |             |             |            |            |      |
|             |                          |             |             |            |            |      |
|             | gcaagatagg               |             |             |            |            |      |
| aggaagtaat  | gagagaaatc               | acagaacccc  | aaggeeatga  | cettereste | ccatggcctt | 4220 |
| aatetteege  | ttcctcgctc               | actgactcgc  | tgegeteggt  | cgcccggccg | cggcgagcgg | 4320 |
| tatcagetea  | ctcaaaggcg               | gtaatacggt  | tateeacaga  | accaggggat | aacgcaggaa | 4380 |
| agaacatgtg  | agcaaaaggc               | cagcaaaagg  | ccaggaaccg  | taaaaaggcc | gegeegeegg | 4440 |
| cgtttttcca  | taggctccgc               | cccctgacg   | agcatcacaa  | aaatcgacgc | tcaagtcaga | 4500 |
| ggtggcgaaa  | cccgacagga               | ctataaagat  | accaggcgtt  | tccccctgga | agctccctcg | 4560 |
| tgcgctctcc  | tgttccgacc               | ctgccgctta  | ccggatacct  | gtccgccttt | ctcccttcgg | 4620 |
| gaagcgtggc  | gctttctcat               | agctcacgct  | gtaggtatct  | cagttcggtg | taggtcgttc | 4680 |
| gctccaagct  | gggctgtgtg               | cacgaacccc  | ccgttcagcc  | cgaccgctgc | gccttatccg | 4740 |
| gtaactatcg  | tcttgagtcc               | aacccggtaa  | gacacgactt  | atcgccactg | gcagcagcca | 4800 |
| ctggtaacag  | gattagcaga               | gcgaggtatg  | taggcggtgc  | tacagagttc | ttgaagtggt | 4860 |
| ggcctaacta  | cggctacact               | agaagaacag  | tatttggtat  | ctgcgctctg | ctgaagccag | 4920 |
| ttaccttcgg  | aaaaagagtt               | ggtagctctt  | gatccggcaa  | acaaaccacc | gctggtagcg | 4980 |
| gtggttttt   | tgtttgcaag               | cagcagatta  | cgcgcagaaa  | aaaaggatct | caagaagatc | 5040 |
| ctttgatctt  | ttctacgggg               | tctgacgctc  | agtggaacga  | aaactcacgt | taagggattt | 5100 |
| tggtcatgag  | attatcaaaa               | aggatettea  | cctagatcct  | tttaaattaa | aaatgaagtt | 5160 |
| ttaaatcaat  | ctaaagtata               | tatgagtaaa  | cttggtctga  | cagttaccaa | tgcttaatca | 5220 |
| gtgaggcacc  | tatctcagcg               | atctgtctat  | ttcgttcatc  | catagttgcc | tgactcgggg | 5280 |
| aaaaaaaaca  | ctgaggtctg               | cctcgtgaag  | aaqqtqttqc  | tgactcatac | caggcctgaa | 5340 |
| tcgccccatc  | atccagccag               | aaaqtqaqqq  | agccacggtt  | gatgagaget | ttgttgtagg | 5400 |
| tagaccagtt  | ggtgattttg               | aacttttgct  | ttaccacaga  | acqqtctqcq | ttgtcgggaa | 5460 |
|             | ctgatccttc               |             |             |            |            |      |
|             | agcgtaatgc               |             |             |            |            |      |
| aaactcatca  | agcatcaaat               | gaaactgcaa  | tttattcata  | tcaggattat | caataccata | 5640 |
|             | agccgtttct               |             |             |            |            |      |
|             | tggtatcggt               |             |             |            |            |      |
|             | tcaaaaataa               |             |             |            |            |      |
|             | ggcaaaagct               |             |             |            |            |      |
| cggcgagaac  | tcaaaatcac               | tacgeaceee  | casaccotta  | treatreate | attocaccta | 5940 |
| acgetegeea  | aatacgcgat               | cogcattaat  | aggagaatta  | cacacacaca | transtars  | 6000 |
| agegagaega  | aacacycyac               | cgccgccaaa  | aggacaatta  | caaacayyaa | gatattette | 6060 |
|             | aacactgcca<br>aatgctgttt |             |             |            |            |      |
|             |                          |             |             |            |            |      |
|             | aaatgcttga               |             |             |            |            |      |
|             | tctgtaacat               |             |             |            |            |      |
|             | ggcttcccat               |             |             |            |            |      |
|             | ttatacccat               |             |             |            |            |      |
|             | tcccgttgaa               |             |             |            |            |      |
| agacagtttt  | attgttcatg               | atgatatatt  | cttatettgt  | gcaatgtaac | accagagatt | 6480 |

```
ttgagacaca acgtggcttt cccccccccc ccattattga agcatttatc agggttattg 6540
 tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg 6600
 cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac 6660
 ctataaaaat aggcgtatca cgaggccctt tcgtc
                                                                   6695
 <210> 114
 <211> 6614
 <212> DNA
 <213> Artificial Sequence
 <223> plasmid pVR1012x/s containing HIV genes
 <400> 114
 tegegegettt eggegatgae ggegaaaace tetgacacat geageteeeg gagaeggtea 60
 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tettatgeat getatactgt ttttggettg gggcctatac acccccgctt cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgeggtge tgttaaeggt ggagggeagt gtagtetgag caqtaetegt tgetgeegeg 1800
egegeeacca gacataatag etgacagact aacagactgt teetttecat gggtetttte 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980
gettetggat ggtgatgate tacaacgtge geggeaacet gtgggtgace gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
```

tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgatcaac tgcaacacca 2340

```
gegecateae ceaggeetge ceeaaggtga gettegaeee cateeceate cactaetgeg 2400
cccccgccgg ctacgccatc ctgaagtgca acaacaagac cttcaacggc accggcccct 2460
gcaacaacgt gagcaccatc cagtgcaccc acggcaccaa gcccgtggtg agcacccagc 2520
tgctgctgaa cggcagcctg gccgaggagg agatcatcat ccgcagcaag aacctgaccg 2580
acaacgtgaa gaccatcatc gtgcacctga acgagagcgt ggagatcaac tgcacccgcc 2640
gateggecca etgeaacate ageegeaceg eetggaacaa gaceetgeag gaggtgggca 2700
agaagetgge egageaette eecaacaagg ecateaagtt egecaageae ageggeggeg 2760
acctggagat caccacccac agcttcaact gccgcggcga gttcttctac tgcaacacca 2820
gcagcetgtt caacagcacc tacaccecca acagcaccga gaacatcacc ggcaccgaga 2880
acagcatcat caccatcccc tgccgcatca agcagatcat caacatgtgg cagggcgtgg 2940
geogegecat gtacgecece eccategagg gcateetgac etgeogeage aacateaceg 3000
geetgetget gaccegegae ggeggeaeeg geatgeaega caccgagate tteegeeeeg 3060
agggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag gtggtggaga 3120
tcaagcccct gggcatcgcc cccaccaagg ccaagcttac cgtgcaggtg cgccagctgc 3180
tgagcggcat cgtgcagcag cagagcaacc tgctgcgcgc catcgaggcc cagcagcaca 3240
tgctgcagct gaccgtgtgg ggcatcaagc agctgcagac ccgcgtgctg gccatcgagc 3300
gctacctgcg cgaccagcag ctcgaggaca tctggaacaa catgacctgg atgcagtggg 3360
accgcgagat cagcaactac accaacacca totaccgcct gctggaggac agccagaacc 3420
agcaggagaa gaacgagcag gacctgctgg ccctggacaa gtggcagaac ctgtggacct 3480
ggttcggcat caccaactgg ctgtggtaca tcaagatctt catcaagatc gtgggcggcc 3540
tgateggeet gegeateate ttegeegtge tgageatetg aacaegtggg atceagatet 3600
gctgtgcctt ctagttgcca gccatctgtt gtttgcccct ccccgtgcc ttccttgacc 3660
ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt 3720
ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat 3780
tgggaagaca atagcaggca tgctggggat gcggtgggct ctatgggtac ccaggtgctg 3840
aagaattgac ccggttcctc ctgggccaga aagaagcagg cacatcccct tctctgtgac 3900
acaccetgte caegeeeetg gttettagtt ceageeeeac teataggaca eteatagete 3960
aggagggete egeetteaat eccaeceget aaagtaettg gageggtete teeeteete 4020
atcageceae caaaccaaac etageeteea agagtgggaa gaaattaaag caagatagge 4080
tattaagtgc agagggagag aaaatgcctc caacatgtga ggaagtaatg agagaaatca 4140
tagaatttta aggccatgat ttaaggccat catggcctta atcttccgct tcctcgctca 4200
ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg 4260
taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 4320
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 4380
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 4440
tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc 4500
tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 4560
gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 4620
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 4680
acceggtaag acacgactta tegecactgg cagcagecae tggtaacagg attagcagag 4740
cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 4800
gaagaacagt atttggtate tgegetetge tgaagecagt tacettegga aaaaqagttg 4860
gtagetettg atceggeaaa caaaceaceg etggtagegg tggttttttt gtttgeaage 4920
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 4980
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 5040
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 5100
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 5160
tetgtetatt tegtteatee atagttgeet gaeteggggg gggggggege tgaggtetge 5220
ctcgtgaaga aggtgttgct gactcatacc aggcctgaat cgccccatca tccagccaga 5280
aagtgaggga gccacggttg atgagagctt tgttgtaggt ggaccagttg gtgattttga 5340
actitigett tgccacggaa cggtctgcgt tgtcgggaag atgcgtgatc tgatccttca 5400
actcagcaaa agttcgattt attcaacaaa gccgccgtcc cgtcaagtca gcgtaatgct 5460
ctgccagtgt tacaaccaat taaccaattc tgattagaaa aactcatcga gcatcaaatg 5520
aaactgcaat ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg 5580
taatgaagga gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc 5640
```

```
tgcgattccg actcgtccaa catcaataca acctattaat ttcccctcgt caaaaataag 5700
gttatcaagt gagaaatcac catgagtgac qactgaatcc qqtqaqaatq qcaaaaqctt 5760
atgcatttct ttccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact 5820
cgcatcaacc aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgatc 5880
gctgttaaaa ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag 5940
cgcatcaaca atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt 6000
cccggggatc gcagtggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat 6060
ggtcggaaga ggcataaatt ccgtcaqcca gtttaqtctg accatctcat ctgtaacatc 6120
attggcaacg ctacctttgc catgtttcag aaacaactct ggcgcatcgg gcttcccata 6180
caatcgatag attgtcgcac ctgattgccc gacattatcg cgaqcccatt tatacccata 6240
taaatcagca tocatgttgg aatttaatcg cggcctcgag caagacgttt cccgttgaat 6300
atggeteata acacceettg tattactgtt tatgtaagea gacagtttta ttqtteatga 6360
tgatatattt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtqqctttc 6420
ccccccccc cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 6480
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 6540
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 6600
gaggeeettt egte
```

<210> 115
<211> 6545
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 115 tegegegettt eggegatgae ggegaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagaqtqc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatqa cqtatqttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acqccccta ttqacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcqtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcetecg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acqqaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeac 1560

| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | taacaataaa | gtatgtgtct | 1620 |
|------------|------------|------------|------------|------------|------------|------|
| gaaaatgagc | gtggagattg | ggctcgcacq | gctgacgcag | atggaagact | taaqqcaqcq | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tactaccaca | 1800 |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | tectttecat | qqqtcttttc | 1860 |
|            |            |            | gatatcgcgg |            |            |      |
| ccatgcgcgt | ggagggcatc | cagcgcaact | ggaagcagtg | gtggatctgg | ggcatcctgg | 1980 |
| gcttctggat | ggtgatgatc | tacaacgtgc | gcggcaacct | gtgggtgacc | gtgtactacg | 2040 |
| gcgtgcccgt | gtggaaggag | gccaagacca | ccctgttctg | caccaacaac | gccaaggcct | 2100 |
| acgacgccga | ggtgcacaac | gtgtgggcca | cccacgcctg | cgtgcccacc | gaccccaacc | 2160 |
| cccaggagat | ggtgctggag | aacgtgaccg | agaacttcaa | catgtgggag | aacqacatqq | 2220 |
| tggagcagat | gcaccaggac | atcatcagcc | tgtgggacca | gagectgaag | ccctqcqtqa | 2280 |
| agctgacccc | cctgtgcgtg | accetgeact | gcagcaaccg | atcgatcaac | tgcaacacca | 2340 |
| gcgccatcac | ccaggcctgc | cccaaggtga | gcttcgaccc | catccccatc | cactactgcg | 2400 |
| ccccgccgg  | ctacgccatc | ctgaagtgca | acaacaagac | cttcaacggc | accadcccct | 2460 |
|            |            |            | acggcaccaa |            |            |      |
|            |            |            | agatcatcat |            |            |      |
| acaacgtgaa | gaccatcatc | gtgcacctga | acgagagcgt | ggagatcaac | tgcacccgcc | 2640 |
|            |            |            | cctggaacaa |            |            |      |
| agaagctggc | cgagcacttc | cccaacaagg | ccatcaagtt | cgccaagcac | agcggcggcg | 2760 |
|            |            |            | gccgcggcga |            |            |      |
|            |            |            | tcaacatgtg |            |            |      |
| tgtacgcccc | ccccatcgag | ggcatcctga | cctgccgcag | caacatcacc | agectactae | 2940 |
| tgacccgcga | cggcggcacc | ggcatgcacg | acaccgagat | cttccqcccc | gaggggggg  | 3000 |
|            |            |            | acaagtacaa |            |            |      |
| tgggcatcgc | cccaccaag  | gccaagctta | ccgtgcaggt | gcgccagctg | ctgagcggca | 3120 |
| tcgtgcagca | gcagagcaac | ctgctgcgcg | ccatcgaggc | ccagcagcac | atgctqcaqc | 3180 |
| tgaccgtgtg | gggcatcaag | cagctgcaga | cccgcgtgct | ggccatcgag | cactacctac | 3240 |
| gcgaccagca | gctcgaggac | atctggaaca | acatgacctg | gatgcagtgg | gaccgcgaga | 3300 |
| tcagcaacta | caccaacacc | atctaccgcc | tgctggagga | cagccagaac | cagcaggaga | 3360 |
| agaacgagca | ggacctgctg | gccctggaca | agtggcagaa | cctgtggacc | tggttcggca | 3420 |
| tcaccaactg | gctgtggtac | atcaagatct | tcatcaagat | cgtgggcggc | ctgatcggcc | 3480 |
| tgcgcatcat | cttcgccgtg | ctgagcatct | gaacacgtgg | gatccagatc | tgctgtgcct | 3540 |
| tctagttgcc | agccatctgt | tgtttgcccc | tcccccgtgc | cttccttgac | cctggaaggt | 3600 |
| gccactccca | ctgtcctttc | ctaataaaat | gaggaaattg | catcgcattg | tctgagtagg | 3660 |
| tgtcattcta | ttctgggggg | tggggtgggg | caggacagca | agggggagga | ttgggaagac | 3720 |
| aatagcaggc | atgctgggga | tgcggtgggc | tctatgggta | cccaggtgct | gaagaattga | 3780 |
| cccggttcct | cctgggccag | aaagaagcag | gcacatcccc | ttctctgtga | cacaccctgt | 3840 |
| ccacgcccct | ggttcttagt | tccagcccca | ctcataggac | actcatagct | caggagggct | 3900 |
|            |            |            | ggagcggtct |            |            |      |
| ccaaaccaaa | cctagcctcc | aagagtggga | agaaattaaa | gcaagatagg | ctattaagtg | 4020 |
|            |            |            | aggaagtaat |            |            |      |
|            |            |            | aatcttccgc |            |            |      |
|            |            |            | tatcagctca |            |            |      |
|            |            |            | agaacatgtg |            |            |      |
|            |            |            | cgtttttcca |            |            |      |
| agcatcacaa | aaatcgacgc | tcaagtcaga | ggtggcgaaa | cccgacagga | ctataaagat | 4380 |
| accaggcgtt | tccccctgga | agctccctcg | tgcgctctcc | tgttccgacc | ctgccgctta | 4440 |
| ccggatacct | gtccgccttt | ctcccttcgg | gaagcgtggc | gctttctcat | agctcacgct | 4500 |
| gtaggtatct | cagttcggtg | taggtcgttc | gctccaagct | gggctgtgtg | cacgaacccc | 4560 |
|            |            |            | gtaactatcg |            |            |      |
| gacacgactt | accgccactg | gcagcagcca | ctggtaacag | gattagcaga | gcgaggtatg | 4680 |
|            |            |            | ggcctaacta |            |            |      |
|            |            |            | ttaccttcgg |            |            |      |
| gateeggeaa | acaaaccacc | gctggtagcg | gtggttttt  | tgtttgcaag | cagcagatta | 4860 |

```
cgcgcagaaa aaaaggatot caagaagato otttgatott ttotacqqqq totqacqoto 4920
agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 4980
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 5040
cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 5100
ttcgttcatc catagttgcc tgactcgggg ggggggggcg ctgaggtctg cctcgtgaag 5160
aaggtgttgc tgactcatac caggectgaa tegececate atecagecag aaagtgaggg 5220
agccacggtt gatgagaget ttgttgtagg tggaccagtt ggtgattttg aacttttgct 5280
ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat ctgatccttc aactcagcaa 5340
aagttegatt tatteaacaa ageegeegte eegteaagte agegtaatge tetgeeagtg 5400
ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa 5460
tttattcata tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg 5520
agaaaactca ccgaggcagt tccataggat ggcaagatcc tggtatcggt ctgcgattcc 5580
gactogtoca acatomatac aacotattaa titoocotog tommaaaataa ggitatomag 5640
tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc 5700
tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 5760
caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 5820
aggacaatta caaacaggaa tcgaatgcaa ccggcgcagg aacactgcca gcgcatcaac 5880
aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 5940
cgcagtggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6000
aggeataaat teegteagee agtttagtet gaceatetea tetgtaacat cattggeaac 6060
gctacctttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6120
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6180
atccatgttg gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6240
aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt 6300
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6360
ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6420
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6480
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6540
tcqtc
```

```
<210> 116
<211> 6626
<212> DNA
<213> Artificial Sequence
<223> plasmid pVR1012x/s containing HIV genes
<400> 116
tegegegettt eggegatgac ggegaaaace teegacacat geageteeeg gagacggea 60
cagettgtct gtaageggat geegggagea gacaageeeg teaggggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
```

| ctccgcccca | ttgacgcaaa | tgggcggtag  | gcgtgtacgg | tgggaggtct  | atataagcag | 900  |
|------------|------------|-------------|------------|-------------|------------|------|
| agctcgttta | gtgaaccgtc | agategeetg  | gagacgccat | ccacactatt  | ttgacctcca | 960  |
| tagaagacac | cgggaccgat | ccagcctccg  | cggccgggaa | cagtacatta  | gaacgcggat | 1020 |
| tccccgtgcc | aagagtgacg | taagtaccgc  | ctatagactc | tataqqcaca  | ccctttggc  | 1080 |
| tcttatgcat | gctatactgt | ttttggcttg  | gggcctatac | acccccactt  | ccttatgcta | 1140 |
| taggtgatgg | tatagettag | cctataggtg  | tgggttattg | accattatto  | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc  | cataacatoo | ctctttacca  | caactatctc | 1260 |
| tattggctat | atqccaatac | tctatccttc  | agagactgac | acqqachchq  | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt  | cacatataca | acaacaccat  | ccccgtgcc  | 1380 |
| cgcagttttt | attaaacata | gcgtgggatc  | tccacacaaa | tetegggtae  | gtgttccgga | 1440 |
| catgggctct | tctccqqtaq | caacaaaact  | tccacatcca | agccctggtc  | ccatgcctcc | 1500 |
| agcggctcat | aatcactcaa | cagctccttg  | ctcctaacag | tagaaaaaaaa | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg  | cacaaggccg | taacaataaa  | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | gacticacaca | actaacacaa | atomageage  | taaggcagcg | 1600 |
| gcagaagaag | atgcaggcag | ctgagttgtt  | gtattctcat | aagaatcaca  | ggtaactccc | 1740 |
| gttacaatac | tattaacaat | agagggagt   | gtagtctgag | cagtactage  | tgctgccgcg | 1/40 |
| cacaccacca | gacataatag | ctgacagact  | aacagactgt | teettteest  | gggtcttttc | 1000 |
| tgcagtcacc | atcatcaaca | catataatca  | gatatege   | cccctctccat | gatatcgcca | 1000 |
| ccatgcgcgt | gaagagata  | caccocaact  | gacaccacta | ataatataa   | ggcatcctgg | 1920 |
| acttctagat | ggtgatgatc | tacaacutuc  | ggaagcagcg | grayaterga  | gtgtactacg | 1980 |
| acatacccat | atagaagaa  | accaacgege  | geggeaacte | grgggrgaee  | gegeaceaeg | 2040 |
| acgacgccga | actacacaac | gtatagacca  | cccegctetg | cgccagegae  | gccaaggcct | 2100 |
| cccaggagat | ggtgcacaac | acatasca    | agaaghtaa  | cgtgeeeaee  | gaccccaacc | 2160 |
| tagagagat  | generagan  | atcatcacco  | tatagasaa  | catgtgggag  | aacgacatgg | 2220 |
| agetgaeeee | cctotacata | accateagee  | cgcgggacca | gageetgaag  | ccctgcgtga | 2280 |
| agacgataca | ccegegegeg | accetgeact  | geageaaceg | accgaccaac  | tgcaacacca | 2340 |
| gegeeaceae | ctaggeetge | cccaaggiga  | gettegaeee | catccccatc  | cactactgcg | 2400 |
| acaacaacab | gagagagata | crgaagrgca  | acaacaagac | cttcaacggc  | accggcccct | 2460 |
| tactactase | gagcaccatc | caguguacee  | acggcaccaa | gcccgtggtg  | agcacccage | 2520 |
| acaacataaa | caccatcatc | geegaggagg  | agaccaccac | ccgcagcaag  | aacctgaccg | 2580 |
| ccaacaacaa | caccaccacc | aggatagga   | togagagege | ggagatcaac  | tgcacccgcc | 2540 |
| acasastast | caccegeaag | agcateegea  | ceggeeeegg | ccaggeette  | tacgccaccg | 2700 |
| agaccetaca | cggcgacacc | cgccaggccc  | accycaacac | cagccgcacc  | gcctggaaca | 2760 |
| tegeceagea | ggaggeggge | aagaageegg  | ccgagcactt | ccccaacaag  | gccatcaagt | 2820 |
| agttetteta | cageggegge | gaccuggaga  | caccaccca  | cagcttcaac  | tgccgcggcg | 2880 |
| ageteeteea | aggegggggg | cgategatee  | cctgccgcat | caagcagatc  | atcaacatgt | 2940 |
| ggcagggcgc | gggcegegee | atgracgece  | ccccatcga  | gggcatcctg  | acctgccgca | 3000 |
| tattagagag | eggeetgetg | cigaeeegeg  | acggeggeac | cggcatgcac  | gacaccgaga | 3060 |
| agatagtaga | cgagggegge | gacatgegeg  | acaactggcg | cagcgagctg  | tacaagtaca | 3120 |
| tacacanact | gateaageee | crgggcarcg  | ccccaccaa  | ggccaagctt  | accgtgcagg | 3180 |
| cccaccage  | getgagegge | ategegeage  | agcagagcaa | cctgctgcgc  | gccatcgagg | 3240 |
|            |            |             |            |             | acccgcgtgc |      |
| ggccaccga  | gegetaeetg | cgcgaccage  | agetegagga | catctggaac  | aacatgacct | 3360 |
| agacgcagcg | ggaccgcgag | accagcaact  | acaccaacac | catctaccgc  | ctgctggagg | 3420 |
| acagetagaa | ctagtaggag | aagaacgage  | aggacctgct | ggccctggac  | aagtggcaga | 3480 |
| tectgeggae | ceggeeegge | atcaccaact  | ggctgtggta | catcaagatc  | ttcatcaaga | 3540 |
| ccgcgggcgg | -bb-b      | ctgcgcatca  | rcttcgccgt | gctgagcatc  | tgaacacgtg | 3600 |
| ggatteatta | ergergree  | trectagetge | cagccatctg | ttgtttgccc  | ctccccgtg  | 3660 |
| cccccccga  | ccctggaagg | cyccaeteee  | actgtcctt  | cctaataaaa  | tgaggaaatt | 3720 |
| gcaccgcacc | gcccgagcag | grgreattet  | attetggggg | araaaaraaa  | gcaggacagc | 3780 |
| eadadadaga | accyggaaga | caacagcagg  | catgetgggg | atgcggtggg  | ctctatgggt | 3840 |
| acceaggige | Lyaagaattg | acceggttee  | tectgggcca | gaaagaagca  | ggcacatccc | 3900 |
| ccccccccc  | acacaccctg | ccacgcccc   | tggttcttag | ttccagcccc  | actcatagga | 3960 |
| cacccatage | ccaggagggc | cccgccttca  | atcccacccg | ctaaagtact  | tggagcggtc | 4020 |
| retecetece | ccaccagccc | accaaaccaa  | acctagcctc | caagagtggg  | aagaaattaa | 4080 |
| agcaagacag | gctattaagt | gcagagggag  | agaaaatgcc | tccaacatgt  | gaggaagtaa | 4140 |

```
tgagagaaat catagaattt taaggccatg atttaaggcc atcatggcct taatcttccg 4200
cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc 4260
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 4320
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 4380
ataggeteeg ecceetgae gageateaca aaaategaeg etcaagteag aggtggegaa 4440
accegacagg actataaaga taccaggegt tteceeetgg aageteeete gtgegetete 4500
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 4560
cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 4620
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 4680
gtettgagte caacceggta agacacgaet tategeeact qqeaqeaqee actqqtaaca 4740
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 4800
acggctacac tagaagaaca gtatttggta totgegetet getgaageca gttacetteg 4860
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 4920
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 4980
tttctacggg gtctgacget cagtggaacg aaaactcacg ttaagggatt ttggtcatga 5040
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 5100
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 5160
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcggg ggggggggc 5220
gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata ccaggcctga atcgcccat 5280
catccagcca gaaagtgagg gagccacggt tgatgagagc tttgttgtag gtggaccagt 5340
tggtgatttt gaacttttgc tttgccacgg aacggtctgc gttgtcggga agatgcgtga 5400
tetgateett caactcagca aaagttegat ttattcaaca aageegeegt eeegtcaagt 5460
cagogtaatg ctctgccagt gttacaacca attaaccaat tctgattaga aaaactcatc 5520
gagcatcaaa tgaaactgca atttattcat atcaggatta tcaataccat atttttgaaa 5580
aagceghthe tghaatgaag gagaaaache acegaggeag thecahagga hggeaagahe 5640
ctggtatcgg tctgcgattc cgactcgtcc aacatcaata caacctatta atttcccctc 5700
gtcaaaaata aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa 5760
tggcaaaagc ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc 5820
atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg 5880
aaatacgcga tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgcag 5940
gaacactgcc agcgcatcaa caatattttc acctgaatca ggatattctt ctaatacctg 6000
gaatgctgtt ttcccgggga tcgcagtggt gagtaaccat gcatcatcag gagtacggat 6060
aaaatgettg atggteggaa gaggeataaa tteegteage eagtttagte tgaccatete 6120
atctgtaaca tcattggcaa cgctaccttt gccatgtttc agaaacaact ctggcgcatc 6180
gggcttccca tacaatcgat agattgtcgc acctgattgc ccgacattat cgcgagccca 6240
tttataccca tataaatcag catccatgtt ggaatttaat cgcggcctcg agcaagacgt 6300
ttcccgttga atatggctca taacacccct tgtattactg tttatgtaag cagacagttt 6360
tattgttcat gatgatatat ttttatcttg tgcaatgtaa catcagagat tttgagacac 6420
aacgtggett teeceeeee eecattattg aageatttat cagggttatt gteteatgag 6480
cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc 6540
ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa 6600
taggogtate acgaggeeet thegte
```

```
<210> 117
<211> 6743
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 117
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
```

WO 02/32943 PCT/US01/25721
Cagettatet ataageggat accagaage gaéaageeg teaggagga teaggagga 120

| caaattatat | ataaaaaat  | 20000000   |            | L          |            |      |
|------------|------------|------------|------------|------------|------------|------|
| ttggggggt  | gtaagcggat | gccgggagea | gacaageeeg | rcagggegeg | ccagegggtg | 120  |
| ceggegggeg | tcggggctgg | cctaactatg | cggcaccaga | gcagattgta | ctgagagtgc | 180  |
| accatatgeg | gtgtgaaata | ccgcacagac | gcgtaaggag | aaaataccgc | accagattgg | 240  |
| ctattggcca | ttgcatacgt | tgtatccata | tcataatatg | tacatttata | ttggctcatg | 300  |
| tecaacatta | ccgccatgtt | gacattgatt | attgactagt | tattaatagt | aatcaattac | 360  |
| ggggtcatta | gttcatagcc | catatatgga | gttccgcgtt | acataactta | cggtaaatgg | 420  |
| cccgcctggc | tgaccgccca | acgacccccg | cccattgacg | tcaataatga | cgtatgttcc | 480  |
| catagtaacg | ccaataggga | ctttccattg | acgtcaatgg | gtggagtatt | tacggtaaac | 540  |
| tgcccacttg | gcagtacatc | aagtgtatca | tatgccaagt | acgcccccta | ttgacgtcaa | 600  |
| tgacggtaaa | tggcccgcct | ggcattatgc | ccagtacatg | accttatggg | actttcctac | 660  |
| ttggcagtac | atctacgtat | tagtcatcgc | tattaccatg | gtgatgcggt | tttggcagta | 720  |
| catcaatggg | cgtggatagc | ggtttgactc | acggggattt | ccaagtctcc | accccattga | 780  |
| cgtcaatggg | agtttgtttt | ggcaccaaaa | tcaacgggac | tttccaaaat | gtcgtaacaa | 840  |
| ctccgcccca | ttgacgcaaa | tgggcggtag | gcgtgtacgg | tgggaggtct | atataagcag | 900  |
| agctcgttta | gtgaaccgtc | agatcgcctg | gagacgccat | ccacgctgtt | ttgacctcca | 960  |
| tagaagacac | cgggaccgat | ccagcctccg | cggccgggaa | cggtgcattg | gaacgcggat | 1020 |
| tccccgtgcc | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca | cccctttqqc | 1080 |
| tettatgeat | gctatactgt | ttttggcttg | gggcctatac | acccccgctt | ccttatecta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg | tgggttattg | accattattq | accactccc  | 1200 |
| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttqcca | caactatete | 1260 |
| tattggctat | atgccaatac | tctgtccttc | agagactgac | acqqactctq | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacgccgt | ccccatacc  | 1380 |
| cgcagttttt | attaaacata | gcgtgggatc | tccacqcqaa | tctcgggtac | atattccaa  | 1440 |
| catgggctct | tctccggtag | caacaaaact | tccacatcca | agccctggtc | ccatacctcc | 1500 |
| agcggctcat | ggtcgctcgg | cageteettg | ctcctaacag | tagaagccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | taacaataaa | atatatatat | 1620 |
| gaaaatgagc | gtggagattg | gactcacaca | actaacacaa | abqqaaqact | taagggaggg | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | agtaactccc | 1740 |
| gttgcggtgc | tgttaacggt | qqaqqqcaqt | gtagtctgag | cagtactcgt | tactaccaca | 1800 |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | tectttecat | agatettte  | 1860 |
| tgcagtcacc | gtcgtcgaca | catataatca | gatatogogg | ccactctaga | gatatogoda | 1920 |
| ccatgcgcgt | ggagggcatc | caqcqcaact | qqaaqcaqtq | gtggatctgg | ggcatcctog | 1980 |
| gcttctggat | gġtgatgatc | tacaacqtqc | gcggcaacct | gtagataacc | gtgtactacg | 2040 |
| gcgtgcccqt | gtggaaggag | gccaagacca | ccctattcta | caccaacaac | accasaacct | 2100 |
| acgacgccga | ggtgcacaac | gtgtgggcca | cccacaccta | catacccacc | gaccccaacc | 2160 |
| cccaqqaqat | ggtgctggag | aacgtgaccg | agaacttcaa | catgtgggag | aacgacatgg | 2220 |
| tggagcagat | gcaccaggac | atcatcagec | totoggacca | gagcctgaag | ccctacatas | 2280 |
| agctgacccc | cctgtgcgtg | accetgeact | acaacaacca | atcgaagaac | tacaacttca | 2340 |
| acatgaccac | cgaggtgcgc | gacaagcgcg | agaaggtgca | caccetatte | taccacctoa | 2400 |
| acatcqtqcc | cctgaagaac | gagagcagca | acaccagegg | cgactaccgc | ctgatcaact | 2460 |
| qcaacaccaq | cgccatcacc | cagacctacc | ccaaggtgag | cttcgacccc | atccccatcc | 2520 |
| actactgcgc | ccccgccggc | tacoccatco | tgaagtgcaa | caacaagacc | ttcaacqqca | 2580 |
| ccqqcccctq | caacaacgtg | agcaccatcc | agtocaccca | caacaccaaa | cccatagtas | 2640 |
| gcacccagct | gctgctgaac | agcagcctag | ccaaggagga | gatcatcatc | cocaocaaga | 2700 |
| acctgaccga | caacgtgaag | accatcatca | tacacctgaa | caagagcata | gagatcaact | 2760 |
| qcacccqccq | atcggcccac | tgcaacatca | accacaccac | ctocaacaac | accetacea  | 2820 |
| aggtggggaa | gaagctggcc | gagcacttcc | ccaacaaggc | catcaagttc | accetgeagg | 2020 |
| qcqqcqqcqa | cctggagatc | accacccaca | acttcaacta | ccacaacasa | ttcttctact | 2000 |
| qcaacaccag | cagcctgttc | aacagcacct | acacccccaa | cadcaccdad | aacatcacco | 3000 |
| gcaccgagaa | cagcatcatc | accateceet | gccgcatcaa | acadatcato | aacatotooc | 3060 |
| agggcqtqqq | ccgcgccatg | tacqccccc  | ccatcgaggg | catectoace | taccacaaca | 3120 |
| acatcaccgg | cctgctgctg | acccqcqacq | dcddcgccaa | catacacaac | accoacatct | 3180 |
| teegeecega | gggcggcgac | atgcgcgaca | actogcocao | cdadctdtac | aagtacaagg | 3240 |
| tggtggagat | caagcccctg | ggcatcgccc | ccaccaagge | caagettace | atacaaatac | 3300 |
| gccagctqct | gagcggcatc | gtgcagcagc | agagcaacch | actacacaca | atcoacccc  | 3360 |
|            |            |            |            | JJ-J-5-0   |            |      |

|             |            |             |            |            | cgcgtgctgg               |      |
|-------------|------------|-------------|------------|------------|--------------------------|------|
|             |            |             |            |            | atgacctgga               |      |
|             |            |             |            |            | ctggaggaca               |      |
|             |            |             |            |            | tggcagaacc               |      |
| tgtggacctg  | gttcggcatc | accaactggc  | tgtggtacat | caagatcttc | atcaagatcg               | 3660 |
| tgggcggcct  | gatcggcctg | cgcatcatct  | tcgccgtgct | gagcatctga | acacgtggga               | 3720 |
| tccagatctg  | ctgtgccttc | tagttgccag  | ccatctgttg | tttgcccctc | ccccgtgcct               | 3780 |
| tecttgacec  | tggaaggtgc | cactcccact  | gtcctttcct | aataaaatga | ggaaattgca               | 3840 |
| tcgcattgtc  | tgagtaggtg | tcattctatt  | ctggggggtg | gggtggggca | ggacagcaag               | 3900 |
| ggggaggatt  | gggaagacaa | tagcaggcat  | gctggggatg | cggtgggctc | tatgggtacc               | 3960 |
|             |            |             |            |            | acatcccctt               |      |
|             |            |             |            |            | cataggacac               |      |
|             |            |             |            |            | agcggtctct               |      |
| ccctccctca  | tcaqcccacc | aaaccaaacc  | tagcctccaa | gagtgggaag | aaattaaagc               | 4200 |
| aagataggct  | attaagtgca | gagggagaga  | aaatgcctcc | aacatotoao | gaagtaatga               | 4260 |
|             |            |             |            |            | tetteegett               |      |
|             |            |             |            |            | tcagctcact               |      |
|             |            |             |            |            | aacatgtgag               |      |
| caaaaggggg  | aceseggeca | aggaaccgta  | sassagacaa | attactacca | tttttccata               | 4500 |
|             |            |             |            |            | tggcgaaacc               |      |
|             |            |             |            |            |                          |      |
|             |            |             |            |            | cgctctcctg<br>agcgtggcgc |      |
|             |            |             |            |            |                          |      |
|             |            |             |            |            | tccaagctgg               |      |
| theochago   | agagatasas | geedageeeg  | accyctycyc | cccaceeggt | aactatcgtc               | 4800 |
| ttagageccaa | cccggcaaga | cacyacttat. | cgccactggc | ageageeaet | ggtaacagga               | 4860 |
| ccagcagage  | gaggiaigia | ggeggegeta  | cagagetett | gaagtggtgg | cctaactacg               | 4920 |
|             |            |             |            |            | accttcggaa               |      |
|             |            |             |            |            | ggttttttg                |      |
|             |            |             |            |            | ttgatctttt               |      |
|             |            |             |            |            | gtcatgagat               |      |
|             |            |             |            |            | aaatcaatct               |      |
|             |            |             |            |            | gaggcaccta               |      |
| ceceagegat  | ctgtctattt | cgttcatcca  | tagttgeetg | accegggggg | ggggggcgct               | 5340 |
|             |            |             |            |            | gccccatcat               |      |
| ccagccagaa  | agcgagggag | ccacggctga  | cgagagettt | gregraggeg | gaccagttgg               | 5460 |
|             |            |             |            |            | tgcgtgatct               |      |
|             |            |             |            |            | gtcaagtcag               |      |
| cgtaatgete  | tgccagtgtt | acaaccaatt  | aaccaattct | gattagaaaa | actcatcgag               | 5640 |
|             |            |             |            |            | tttgaaaaag               |      |
|             |            |             |            |            | caagatcctg               |      |
|             |            |             |            |            | teceetegte               |      |
| aaaaataagg  | ttatcaagtg | agaaatcacc  | atgagtgacg | actgaatccg | gtgagaatgg               | 5880 |
| caaaagctta  | tgcatttctt | tccagacttg  | ttcaacaggc | cagccattac | gctcgtcatc               | 5940 |
|             |            |             |            |            | cgagacgaaa               |      |
| tacgcgatcg  | ctgttaaaag | gacaattaca  | aacaggaatc | gaatgcaacc | ggcgcaggaa               | 6060 |
| cactgccagc  | gcatcaacaa | tattttcacc  | tgaatcagga | tattcttcta | atacctggaa               | 6120 |
| tgctgttttc  | ccggggatcg | cagtggtgag  | taaccatgca | tcatcaggag | tacggataaa               | 6180 |
| atgcttgatg  | gtcggaagag | gcataaattc  | cgtcagccag | tttagtctga | ccatctcatc               | 6240 |
| tgtaacatca  | ttggcaacgc | tacctttgcc  | atgtttcaga | aacaactctg | gcgcatcggg               | 6300 |
| cttcccatac  | aatcgataga | ttgtcgcacc  | tgattgcccg | acattatcgc | gagcccattt               | 6360 |
| atacccatat  | aaatcagcat | ccatgttgga  | atttaatcgc | ggcctcgagc | aagacgtttc               | 6420 |
| ccgttgaata  | tggctcataa | caccccttgt  | attactgttt | atgtaagcag | acagttttat               | 6480 |
| tgttcatgat  | gatatattt  | tatcttgtgc  | aatgtaacat | cagagatttt | gagacacaac               | 6540 |
| gtggctttcc  | cccccccc   | attattgaag  | catttatcag | ggttattgtc | tcatgagcgg               | 6600 |
| atacatattt  | gaatgtattt | agaaaaataa  | acaaataggg | gttccgcgca | catttccccg               | 6660 |

aaaagtgcca cctgacgtct aagaaaccat tattatcatg acattaacct ataaaaatag 6720 gcgtatcacg aggccctttc gtc 6743

```
<210> 118
<211> 6674
<212> DNA
<213> Artificial Sequence
<223> plasmid pVR1012x/s containing HIV genes
<400> 118
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgq 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cqtatqttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaaccgtc agategeetg gagacgeeat ceaegetgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eccetttgge 1080
tottatgcat gotatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacqccqt cccccqtqcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taagqcaqcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctqccgcq 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980
gettetggat ggtgatgate tacaacgtge geggeaacet gtgggtgace gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgaagaac tgcagcttca 2340
acatgaccac cgaggtgcgc gacaagcgcg agaaggtgca cgccctgttc taccgcctgg 2400
acategtgee cetgaagaac gagageagea acaceagegg egactacege etgateaact 2460
```

|   | gcaacaccag  | cgccatcacc | caggcctgcc  | ccaaggtgag  | cttcgacccc  | atccccatcc | 2520 |
|---|-------------|------------|-------------|-------------|-------------|------------|------|
| • | actactgcgc  | ccccgccggc | tacgccatcc  | tgaagtgcaa  | caacaagacc  | ttcaacggca | 2580 |
|   | ccggcccctg  | caacaacgtg | agcaccatcc  | agtgcaccca  | cggcaccaag  | cccgtggtga | 2640 |
|   | gcacccagct  | gctgctgaac | ggcagcctgg  | ccgaggagga  | gatcatcatc  | cgcagcaaga | 2700 |
|   | acctgaccga  | caacgtgaag | accatcatcg  | tgcacctgaa  | cgagagcgtg  | gagatcaact | 2760 |
|   | gcacccgccg  | atcggcccac | tgcaacatca  | gccgcaccgc  | ctggaacaag  | accetgeagg | 2820 |
|   | aggtgggcaa  | gaagctggcc | gagcacttcc  | ccaacaaggc  | catcaagttc  | gccaagcaca | 2880 |
|   | gcggcggcga  | cctggagatc | accacccaca  | gcttcaactg  | ccqcqqcqaq  | ttcttctact | 2940 |
|   | gcaacacccg  | atcgatcccc | tgccgcatca  | agcagatcat  | caacatgtgg  | cagagcatag | 3000 |
|   | gccgcgccat  | gtacgccccc | cccatcgagg  | gcatcctgac  | ctqccqcaqc  | aacatcacco | 3060 |
|   | gcctgctgct  | gacccgcgac | ggcggcaccg  | gcatgcacga  | caccgagate  | ttccaccca  | 3120 |
|   | agggcggcga  | catgcgcgac | aactggcgca  | gcgagctgta  | caaqtacaaq  | ataataaaa  | 3180 |
|   | tcaagcccct  | gggcatcgcc | cccaccaagg  | ccaagettae  | catacagata  | caccaactac | 3240 |
|   | tgagcggcat  | cgtgcagcag | cagagcaacc  | tactacacac  | catcgaggcc  | cagcagcaga | 3300 |
|   | tactacaact  | gaccgtgtgg | ggcatcaagc  | agctgcagac  | ccacatacta  | accatcaaca | 3360 |
|   | gctacctgcg  | cgaccagcag | ctcgaggaca  | tetagaacaa  | catgacctcg  | atacagtage | 3420 |
|   | accocoagat  | cagcaactac | accaacacca  | tetaccacct  | actagaccegg | acgcagcggg | 3400 |
|   | aggaggagaa  | gaacgagcag | acctactaa   | ccctaccacca | gttggaggat  | agecagaace | 3460 |
|   | agttcaccat  | caccaactgg | ctataataca  | tcaagatott  | geggeagaac  | ctgtggatet | 3540 |
|   | tastcaacet  | acceatcata | ttggggtaca  | tanagatete  | cattaagatt  | gragaeagee | 3600 |
|   | actatacett  | gcgcatcatc | geestatatt  | atttagagat  | aacacgtggg  | acccagatet | 3660 |
|   | chagagagata | ctagttgcca | tataatttaa  | tantanata   | eccegugee   | tteettgace | 3/20 |
|   | ctggaaggtg  | ccactcccac | tatagagagat | caacaaaacg  | aggaaattge  | accgcattgt | 3780 |
|   | tagageagge  | gtcattctat | terggggggt  | 9999199990  | aggacagcaa  | gggggaggat | 3840 |
|   | aggaagaca   | atagcaggca | cgccggggat  | geggeggee   | ctatgggtac  | ccaggtgctg | 3900 |
|   | aagaactgac  | ccggttcctc | etgggecaga  | aagaagcagg  | cacateceet  | tctctgtgac | 3960 |
|   |             | cacgcccctg |             |             |             |            |      |
|   | aggagggete  | cgccttcaat | eccaeceget  | aaagtacttg  | gagcggtctc  | tecetecete | 4080 |
|   | accageceae  | caaaccaaac | ctagceteca  | agagtgggaa  | gaaattaaag  | caagataggc | 4140 |
|   | tattaagtgc  | agaggagag  | aaaatgcctc  | caacatgtga  | ggaagtaatg  | agagaaatca | 4200 |
|   | cagaatttta  | aggccatgat | ttaaggccat  | catggcctta  | atcttccgct  | tectegetea | 4260 |
|   | ctgactcgct  | gcgctcggtc | gttcggctgc  | ggcgagcggt  | atcageteae  | tcaaaggcgg | 4320 |
|   | taatacggtt  | atccacagaa | tcaggggata  | acgcaggaaa  | gaacatgtga  | gcaaaaggcc | 4380 |
|   | agcaaaaggc  | caggaaccgt | aaaaaggccg  | cgttgctggc  | gtttttccat  | aggctccgcc | 4440 |
|   | cccctgacga  | gcatcacaaa | aatcgacgct  | caagtcagag  | gtggcgaaac  | ccgacaggac | 4500 |
|   | tataaagata  | ccaggcgttt | ccccctggaa  | gctccctcgt  | gcgctctcct  | gttccgaccc | 4560 |
|   | tgccgcttac  | cggatacctg | teegeettte  | tcccttcggg  | aagcgtggcg  | ctttctcata | 4620 |
|   | gctcacgctg  | taggtatctc | agttcggtgt  | aggtcgttcg  | ctccaagctg  | ggctgtgtgc | 4680 |
|   | acgaaccccc  | cgttcagccc | gaccgctgcg  | ccttatccgg  | taactatcgt  | cttgagtcca | 4740 |
|   | acccggtaag  | acacgactta | tegecaetgg  | cagcagccac  | tggtaacagg  | attagcagag | 4800 |
|   | cgaggtatgt  | aggcggtgct | acagagttct  | tgaagtggtg  | gcctaactac  | ggctacacta | 4860 |
|   | gaagaacagt  | atttggtatc | tgcgctctgc  | tgaagccagt  | taccttcgga  | aaaagagttg | 4920 |
|   | gtagctcttg  | atccggcaaa | caaaccaccg  | ctggtagcgg  | tggtttttt   | gtttgcaagc | 4980 |
|   | agcagattac  | gcgcagaaaa | aaaggatctc  | aagaagatcc  | tttgatcttt  | tctacggggt | 5040 |
|   | ctgacgctca  | gtggaacgaa | aactcacgtt  | aagggatttt  | ggtcatgaga  | ttatcaaaaa | 5100 |
|   | ggatcttcac  | ctagatectt | ttaaattaaa  | aatgaagttt  | taaatcaatc  | taaagtatat | 5160 |
|   | atgagtaaac  | ttggtctgac | agttaccaat  | gcttaatcag  | tgaggcacct  | atctcagcga | 5220 |
|   | tctgtctatt  | tcgttcatcc | atagttgcct  | gactcggggg  | gggggggcgc  | tgaggtctgc | 5280 |
|   | ctcgtgaaga  | aggtgttgct | gactcatacc  | aggcctgaat  | cgccccatca  | tccagccaga | 5340 |
|   | aagtgaggga  | gccacggttg | atgagagctt  | tgttgtaggt  | ggaccagttg  | gtgattttga | 5400 |
|   | acttttgctt  | tgccacggaa | cggtctgcgt  | tgtcgggaag  | atgcgtgatc  | tgatccttca | 5460 |
|   | actcagcaaa  | agttcgattt | attcaacaaa  | gccgccgtcc  | cgtcaagtca  | gcgtaatgct | 5520 |
|   | ctgccagtgt  | tacaaccaat | taaccaattc  | tgattagaaa  | aactcatcaa  | gcatcaaato | 5580 |
|   | aaactgcaat  | ttattcatat | caggattatc  | aataccatat  | ttttqaaaaa  | gccqtttcta | 5640 |
|   | taatgaagga  | gaaaactcac | cgaggcagtt  | ccataggatg  | gcaagatcct  | ggtatcootc | 5700 |
|   | tgcgattccg  | actcgtccaa | catcaataca  | acctattaat  | ttcccctcat  | caaaaataaa | 5760 |
|   | _           | •          |             |             |             |            |      |

```
gttatcaagt gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt 5820
atgeatttet ttecagaett gtteaacagg ceagceatta egetegteat caaaateact 5880
egeateaace aaacegttat teattegtga ttgegeetga gegagaegaa ataegegate 5940
gctgttaaaa ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag 6000
cgcatcaaca atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt 6060
cccggggatc gcagtggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat 6120
ggtcggaaga ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc 6180
attggcaacg ctacctttgc catgtttcag aaacaactct ggcgcatcgg gcttcccata 6240
caatcgatag attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata 6300
taaatcaqca tccatqttqq aatttaatcq cqqcctcqaq caaqacqttt cccqttqaat 6360
atggctcata acacccttg tattactgtt tatgtaagca gacagtttta ttgttcatga 6420°
tgatatattt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc 6480
cccccccc cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 6540
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 6600
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 6660
gaggcccttt cgtc
```

<210> 119 <211> 6755 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 119 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660

ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtetec accecattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcetecg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtacege etatagaete tataggeaca eccetttgge 1080 tottatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 egeagttttt attaaacata gegtgggate tecaegegaa tetegggtae gtgtteegga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ecatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaaeag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620

| ~~~~~~~    | ~+~~~~     |             |             |            |            |      |
|------------|------------|-------------|-------------|------------|------------|------|
| gaaaatgagc | gragatra   | ggctcgcacg  | gctgacgcag  | atggaagact | taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt  | gtattctgat  | aagagtcaga | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt | ggagggcagt  | gtagtctgag  | cagtactcgt | tgctgccgcg | 1800 |
| cgcgccacca | gacataatag | ctgacagact  | aacagactgt  | tcctttccat | gggtcttttc | 1860 |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca  | gatatcgcgg  | ccgctctaga | gatatcgcca | 1920 |
| ccatgcgcgt | ggagggcatc | cagcgcaact  | ggaagcagtg  | gtggatctgg | ggcatcctgg | 1980 |
| gcttctggat | ggtgatgatc | tacaacgtgc  | gcggcaacct  | gtgggtgacc | gtgtactacg | 2040 |
| gcgtgcccgt | gtggaaggag | gccaagacca  | ccctgttctg  | cgccagcgac | gccaaggcct | 2100 |
| acgacgccga | ggtgcacaac | gtgtgggcca  | cccacgcctg  | cgtgcccacc | gaccccaacc | 2160 |
| cccaggagat | ggtgctggag | aacgtgaccg  | agaacttcaa  | catgtgggag | aacgacatgg | 2220 |
| tggagcagat | gcaccaggac | atcatcagcc  | tgtgggacca  | gagcctgaag | ccctgcgtga | 2280 |
| agctgacccc | cctgtgcgtg | accctgcact  | gcagcaaccg  | atcgaagaac | tgcagcttca | 2340 |
| acatgaccac | cgaggtgcgc | gacaagcgcg  | agaaggtgca  | cgccctgttc | taccgcctgg | 2400 |
| acatcgtgcc | cctgaagaac | gagagcagca  | acaccagegg  | cgactaccgc | ctgatcaact | 2460 |
| gcaacaccag | cgccatcacc | caggeetgee  | ccaaggtgag  | cttcgacccc | atccccatcc | 2520 |
| actactgcgc | ccccqccqqc | tacqccatcc  | tgaagtgcaa  | caacaagacc | ttcaacggca | 2580 |
| ccggcccctg | caacaacgtg | agcaccatcc  | agtgcaccca  | CGGCaccaag | cccgtggtga | 2640 |
| gcacccagct | qctqctqaac | ggcagcctgg  | ccgaggagga  | gatcatcatc | cgcagcaaga | 2700 |
| acctgaccga | caacqtqaaq | accatcatca  | tgcacctgaa  | caagagcata | gagatcaact | 2760 |
| qcacccqccc | caacaacaac | acccgcaaga  | gcatccgcat  | caaccccaac | caggccttct | 2820 |
| acqccaccqq | cgagatcatc | ggcgacatcc  | gccaggccca  | ctgcaacatc | agccgcaccg | 2880 |
| cctqqaacaa | gaccctgcag | gaggtgggca  | agaagctggc  | cgaggacttc | cccaacaagg | 2940 |
| ccatcaaqtt | cqccaaqcac | agcagcagca  | acctggagat  | caccacccac | agcttcaact | 3000 |
| accacaacaa | gttcttctac | tgcaacaccc  | gatcgatccc  | ctaccacate | aagcagatca | 3060 |
| tcaacatoto | acaaaacata | ggccgcgcca  | tatacacccc  | ccccatcaaa | ggcatcctga | 3120 |
| cctaccacaa | caacatcacc | gacctactac  | tgacccgcga  | caacaacac  | ggcatgcacg | 3780 |
| acaccgagat | cttccacccc | gaggggggg   | acatococoa  | caactggcgc | agcgagctgt | 3240 |
| acaagtacaa | gataataaaa | atcaagcccc  | tagacatcac  | ccccaccaan | gccaagctta | 3270 |
| ccatacaaat | acaccaacta | ctgagcggca  | tegtgeagea  | acadadcaac | ctgctgcgcg | 3360 |
| ccatcgagge | ccagcagcac | atactacaac  | taaccatata  | gaacatcaaa | cagctgcaga | 3420 |
| cccacatact | gaccatcgag | cactacctac  | acasccaaca  | actcasaas  | atctggaaca | 3480 |
| acatgacctg | gatgcagtgg | gaccgcgaga  | tcagcaacta  | caccaacacc | atctaccgcc | 3540 |
| tactagagga | cagccagaac | cagcaggaga  | agaacgagca  | ggacctgctg | gccctggaca | 3600 |
| agtagcagaa | cctatagacc | tagttcagca  | tcaccaactq  | actataatac | atcaagatct | 3660 |
| tcatcaagat | catagacaac | ctgatcggcc  | tacacatcat  | cttcaccata | ctgagcatct | 3720 |
| gaacacgtgg | gatccagatc | tactatacct  | tctagttgcc  | agccatctgt | tgtttgcccc | 3720 |
| tececeatae | cttccttgac | cctggaaggt  | occactocca  | ctatacttta | ctaataaaat | 3040 |
| gaggaaattg | catcocatto | tctgagtagg  | tatcattcta  | ttctagagag | tggggtgggg | 3040 |
| caggacagca | aggggaaga  | ttaggaagac  | aatagcaggc  | atoctoggg  | tgcggtgggc | 3960 |
| tctatgggta | cccaggtact | gaagaattga  | cccaattcct  | cctaaaccaa | aaagaagcag | 4020 |
| gcacatcccc | ttctctgtga | cacaccctgt  | ccacacacac  | agttettagt | tccagcccca | 4020 |
| ctcataggac | actcataget | caggagggct  | ccaccttcaa  | teccaceege | taaagtactt | 4140 |
| ggagcggtct | ctccctccct | catcagecea  | ccaaaccaaa  | cctacctcc  | aagagtggga | 4200 |
| agaaattaaa | gcaagatagg | ctattaagtg  | cadadddada  | gaaaatgggt | ccaacatgtg | 4200 |
| aggaagtaat | gagagaaatc | atacaatttt  | aaddccatda  | tttaaggcca | tcatggcctt | 4200 |
| aatcttccgc | tteeteacte | actgactcgc  | tacactcaat  | cattcaacta | cggcgagcgg | 4300 |
| tatcagetea | ctcaaaggcg | gtaatacggt  | tatccacage  | atcaggggat | aacgcaggaa | 4360 |
| agaacatgtg | agcaaaaagc | Cadcaaaadd  | ccaddaaccd  | taaaaaaaaa | gcgttgctgg | 4500 |
| cqtttttcca | taggchecoe | cccctaaca   | aggatgaga   | aaatcoaccc | tcaagtcaga | 45EN |
| gatagcasss | cccacacaca | ctataaagat  | accadacath  | tececetées | agctccctcg | 4520 |
| tacachetee | tattccaacc | ctaccactta  | ccddatacct  | atcoacett  | ctcccttcgg | 4020 |
| gaagcotooc | actttctcat | ageteacgeta | graggaracct | canttenete | taggtcgttc | 4000 |
| actccaaact | gaactatata | cacqaacccc  | coatteace   | cascccctcc | gccttatccg | 4/40 |
| gtaachatco | tettgagtee | aacccontaa  | gacaccactt  | atrocoscte | gcagcagcca | 4000 |
| ctggtaacag | gattagrage | acasaatsta  | taggcggtgc  | tacacachte | ttgaagtggt | 4000 |
|            | Januageaga |             |             | Lavagagete | ccyaaycyyc | せコムリ |

```
ggcctaacta cggctacact agaagaacag tatttggtat ctqcqctctg ctgaaqccag 4980
ttaccttegg aaaaagagtt ggtagetett gateeggeaa acaaaceace getggtageg 5040
gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 5100
ctttgatett ttctacgggg tctgacgctc agtggaacqa aaactcacqt taaqqqattt 5160
tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 5220
ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca 5280
gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactcqqqq 5340
ggggggggg ctgaggtctg cctcgtgaag aaggtgttgc tgactcatac caggcctgaa 5400
tegececate atccagecag aaagtgaggg agecaeggtt gatgagaget ttgttgtagg 5460
tggaccagtt ggtgattttg aacttttgct ttgccacqqa acqqtctqcq ttqtcqqqaa 5520
gatgcgtgat ctgatccttc aactcagcaa aagttcgatt tattcaacaa agccgccgtc 5580
ccgtcaagtc agcgtaatgc tctgccagtg ttacaaccaa ttaaccaatt ctgattagaa 5640
aaactcatcg agcatcaaat gaaactgcaa tttattcata tcaggattat caataccata 5700
tttttgaaaa agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat 5760
ggcaagatcc tggtatcggt ctgcgattcc gactcgtcca acatcaatac aacctattaa 5820
tttcccctcg tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc 5880
cggtgagaat ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt 5940
acgctcgtca tcaaaatcac tcgcatcaac caaaccgtta ttcattcgtg attgcgcctg 6000
agcgagacga aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa 6060
ccggcgcagg aacactgcca gcgcatcaac aatattttca cctgaatcag gatattcttc 6120
taatacctgg aatgctgttt tcccggggat cgcagtggtg agtaaccatg catcatcagg 6180
agtacggata aaatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct 6240
gaccatetea tetgtaacat cattggcaac getacetttg ceatgtttea gaaacaacte 6300
tggcgcatcg ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc 6360
gegageeeat ttataceeat ataaateage atecatgttg gaatttaate geggeetega 6420
gcaagacgtt teeegttgaa tatggeteat aacaceeett gtattaetgt ttatgtaage 6480
agacagtttt attgttcatg atgatatatt tttatcttgt gcaatgtaac atcagagatt 6540
ttgagacaca acgtggcttt ccccccccc ccattattga agcatttatc agggttattg 6600
tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg 6660
cacatttecc cgaaaagtgc cacetgacgt ctaagaaacc attattatca tgacattaac 6720
ctataaaaat aggcgtatca cgaggccctt tcgtc
                                                                  6755
```

```
<210> 120
<211> 6677
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 120
```

| WO 02/32943 | 3          |            |            |            | PCT/US       | 01/2572 |
|-------------|------------|------------|------------|------------|--------------|---------|
| catcaatggg  | cgtggatagc | ggtttgactc | acggggattt | ccaagtctcc | accccattga   | 780     |
|             | agtttgtttt |            |            |            |              |         |
|             | ttgacgcaaa |            |            |            |              |         |
|             | gtgaaccgtc |            |            |            |              |         |
|             | cgggaccgat |            |            |            |              |         |
| tecceqtqce  | aagagtgacg | taagtaccqc | ctatagactc | tataggcaca | cccctttggc   | 1080    |
| tettatgeat  | gctatactgt | ttttggcttg | gggcctatac | acccccgctt | ccttatgcta   | 1140    |
|             | tatagcttag |            |            |            |              |         |
|             | gatactttcc |            |            |            |              |         |
|             | atgccaatac |            |            |            |              |         |
|             | ccatttatta |            |            |            |              |         |
|             | attaaacata |            |            |            |              |         |
|             | tctccggtag |            |            |            |              |         |
|             | ggtcgctcgg |            |            |            |              |         |
|             | ccaccaccac |            |            |            |              |         |
|             | gtggagattg |            |            |            |              |         |
| gcagaagaag  | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc   | 1740    |
| attacaatac  | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg   | 1800    |
|             | gacataatag |            |            |            |              |         |
| tgcagtcacc  | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatcgcca   | 1920    |
| ccatgcgcgt  | ggagggcatc | cagcgcaact | ggaagcagtg | gtggatctgg | ggcatcctgg   | 1980    |
| gcttctggat  | ggtgatgatc | tacaacgtgc | gcggcaacct | gtgggtgacc | gtgtactacg   | 2040    |
| gcgtgcccgt  | gtggaaggag | gccaagacca | ccctgttctg | cgccagcgac | gccaaggcct   | 2100    |
|             | ggtgcacaac |            |            |            |              |         |
| cccaggagat  | ggtgctggag | aacgtgaccg | agaacttcaa | catgtgggag | aacgacatgg   | 2220    |
| tggagcagat  | gcaccaggac | atcatcagcc | tgtgggacca | gagcctgaag | ccctgcgtga   | 2280    |
| agctgacccc  | cctgtgcgtg | accetgeact | gcagcaaccg | caccatcgac | tacaacaacc   | 2340    |
| gcaccgacaa  | catgggcggc | gagatcaaga | actgcagctt | ccgatcgatc | aactgcaaca   | 2400    |
| ccagcgccat  | cacccaggcc | tgccccaagg | tgagcttcga | ccccatcccc | atccactact   | 2460    |
| gegeeeeege  | cggctacgcc | atcctgaagt | gcaacaacaa | gaccttcaac | ggcaccggcc   | 2520    |
| cctgcaacaa  | cgtgagcacc | atccagtgca | cccacggcac | caagcccgtg | gtgagcaccc   | 2580    |
| agctgctgct  | gaacggcagc | ctggccgagg | aggagatcat | cateegeage | aagaacctga   | 2640    |
| ccgacaacgt  | gaagaccatc | atcgtgcacc | tgaacgagag | cgtggagatc | aactgcaccc   | 2700    |
|             | ccactgcaac |            |            |            |              |         |
| gcaagaagct  | ggccgagcac | ttccccaaca | aggccatcaa | gttcgccaag | cacagcggcg   | 2820    |
| gcgacctgga  | gatcaccacc | cacagettea | actgccgcgg | cgagttcttc | tactgcaaca   | 2880    |
| ccagcagcct  | gttcaacagc | acctacaccc | ccaacagcac | cgagaacatc | accggcaccg   | 2940    |
| agaacagcat  | catcaccatc | ccctgccgca | tcaagcagat | catcaacatg | tggcagggcg   | 3000    |
| tgggccgcgc  | catgtacgcc | cccccatcg  | agggcatcct | gacctgccgc | agcaacatca   | 3060    |
|             | gctgacccgc |            |            |            |              |         |
|             | cgacatgcgc |            |            |            |              |         |
| agatcaagcc  | cctgggcatc | gcccccacca | aggccaagct | taccgtgcag | gtgcgccagc   | 3240    |
| tgctgagcgg  | catcgtgcag | cagcagagca | acctgctgcg | cgccatcgag | gcccagcagc   | 3300    |
| acatgctgca  | gctgaccgtg | tggggcatca | agcagctgca | gacccgcgtg | ctggccatcg   | 3360    |
| agcgctacct  | gcgcgaccag | cagetegagg | acatctggaa | caacatgacc | tggatgcagt   | 3420    |
| gggaccgcga  | gatcagcaac | tacaccaaca | ccatctaccg | cctgctggag | gacagccaga   | 3480    |
| accagcagga  | gaagaacgag | caggacctgc | tggccctgga | caagtggcag | aacetgtgga   | 3540    |
| cctggttcgg  | catcaccaac | tggctgtggt | acatcaagat | cttcatcaag | accgtgggcg   | 3600    |
| gcctgatcgg  | cctgcgcatc | atcttcgccg | tgctgagcat | ctgaacacgt | gggatccaga   | 3660    |
|             | cttctagttg |            |            |            |              |         |
| accctggaag  | gtgccactcc | cactgtectt | coccaacaaa | acgaggaaat | cyca ccyca c | 3/00    |
| tgtctgagta  | ggtgtcattc | tattctgggg | aaraaaaraa | ggcaggacag | caagggggag   | 3040    |
| gattgggaag  | acaatagcag | gcatgctggg | gatgeggtgg | gctctatggg | cacccaggcg   | 3300    |
|             | gacccggttc |            |            |            |              |         |
| gacacaccct  | gtccacgccc | ctggttctta | gttccagccc | cactcatagg | acacccatag   | 4020    |

```
ctcaggaggg ctccgccttc aatcccaccc gctaaagtac ttggagcggt ctctccctcc 4080
ctcatcagcc caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata 4140
ggctattaag tgcagaggga gagaaaatgc ctccaacatg tgaggaagta atgagagaaa 4200
tcatagaatt ttaaggccat gatttaaggc catcatggcc ttaatcttcc gcttcctcgc 4260
teactgacte getgegeteg gtegttegge tgeggegage ggtateaget cacteaaagg 4320
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 4380
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 4440
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 4500
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 4560
coolgooget tacoggatac olytocgoot thotocotto qqqaaqcqtq qcqctttctc 4620
atageteaeg etgtaggtat eteagttegg tgtaggtegt tegeteeaag etgggetgtg 4680
tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cqqtaactat cqtcttqaqt 4740
ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 4800
gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 4860
ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 4920
ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca 4980
agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 5040
ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 5100
aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta 5160
tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag 5220
tgcctcgtga agaaggtgtt gctgactcat accaggcctg aatcgcccca tcatccagcc 5340
agaaagtgag ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt 5400
tgaacttttg ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatcct 5460
tcaactcagc aaaagttcga tttattcaac aaagccgccg tcccgtcaag tcagcgtaat 5520
gctctgccag tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa 5580
atgaaactgc aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt 5640
ctgtaatgaa ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg 5700
gtctgcgatt ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat 5760
aaggttatca agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag 5820
cttatgcatt tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc 5880
actegcatea accaaacegt tatteatteg tgattgegee tgagegagae gaaataegeg 5940
ategetgtta aaaggacaat tacaaacagg aategaatge aaceggegea ggaacactge 6000
cagegeatea acaatatttt cacetgaate aggatattet tetaatacet ggaatgetgt 6060
tttcccgggg atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaatgctt 6120
gatggtcgga agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac 6180
atcattggca acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc 6240
atacaatcga tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc 6300
atataaatca gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg 6360
aatatggctc ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca 6420
tgatgatata tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct 6480
ttccccccc ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat 6540
atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccqaaaaqt 6600
gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat 6660
cacgaggccc tttcgtc
```

```
<210> 121
<211> 6677
<212> DNA
<213> Artificial Sequence
```

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 121

tegegegettt eggegatgae ggegaaaaee teegacaeat geageteeeg gagaeggeea 60 cagettgtet gtaageggat geegggagea qaeaageeeq teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cocgcetgge tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg geagtacate aagtgtatea tatgccaagt aegeceeeta ttgaegteaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt tittggettg gggeetatae acceegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae aeggaetetg tatttttaea 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgeggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980 gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040 gegtgeeegt gtggaaggag geeaagaeea ceetgttetg egeeagegae geeaaggeet 2100 acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160 cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220 tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agetgacece cetgtgegtg accetgeact geageaaceg caceategae tacaacaace 2340 gcaccgacaa catgggcggc gagatcaaga actgcagctt ccgatcgatc aactgcaaca 2400 ccagegccat cacccaggcc tgccccaagg tgagettega ccccatecec atecactact 2460 gcgccccgc cggctacgcc atcctgaagt gcaacaacaa gaccttcaac ggcaccggcc 2520 cctgcaacaa cgtgagcacc atccagtgca cccacggcac caagcccgtg qtgagcaccc 2580 agetgetget gaacggeage etggeegagg aggagateat cateegeage aagaacetga 2640 ccgacaacgt gaagaccatc atcgtgcacc tgaacgagag cgtggagatc aactgcaccc 2700 geegategge ceaetgeaac ateageegea cegeetggaa caagaceetg caggaggtgg 2760 gcaagaaget ggccgagcac ttccccaaca aggccatcaa gttcgccaag cacagcggcg 2820 gcgacctgga gatcaccacc cacagcttca actgccgcgg cgagttcttc tactgcaaca 2880 ccagcagect gttcaacage acctacacce ccaacageac egagaacate accggcaceg 2940 agaacagcat catcaccatc ccctgccgca tcaagcagat catcaacatg tggcagggcg 3000 tgggccgcgc catgtacgcc cccccatcg agggcatcct gacctgccgc agcaacatca 3060 ceggeetget getgaceege gaeggeggea eeggeatgea egacacegag atetteegee 3120 ccgagggcgg cgacatgcgc gacaactggc gcagcgagct gtacaagtac aaggtggtgg 3180

| agatcaagcc | cctagacatc | gcccccacca  | aggccaagct  | taccotocao | atacaccaac | 3240 |
|------------|------------|-------------|-------------|------------|------------|------|
|            |            |             | acctgctgcg  |            |            |      |
| acatoctoca | actascata  | tagaagagaa  | agcagctgca  | gecaccgag  | stageage   | 3360 |
| agggetaget | accaaccaa  | cagetcgagg  | acatctggaa  | gaccegegeg | taggecateg | 2720 |
| agagacacac | gegegaceag | tacaccaaca  | ccatctaggaa | catcatgacc | aggacgcagt | 2400 |
| 3994009094 | gaccagcaac | cacaccaaca  | ccatctaccg  | cetgetggag | gacagecaga | 3480 |
|            |            |             | tggccctgga  |            |            |      |
|            |            |             | acatcaagat  |            |            |      |
| babaahahaa | ectgegeate | accttegeeg  | tgctgagcat  | ctgaacacgt | gggatccaga | 3660 |
| tergergrae | cttctagttg | ccagccaccc  | gttgtttgcc  | cccccccdt  | geetteettg | 3720 |
| accetggaag | grgecaetee | cactgtcctt  | tcctaataaa  | atgaggaaat | tgcatcgcat | 3780 |
| tgtetgagta | ggtgtcatte | cattetgggg  | aaraaaaraa  | ggcaggacag | caagggggag | 3840 |
| gattgggaag | acaacagcag | gcatgetggg  | gatgcggtgg  | gctctatggg | tacccaggtg | 3900 |
| ccgaagaacc | gacccggttc | ctcctgggcc  | agaaagaagc  | aggcacatcc | cettetetgt | 3960 |
| gacacaccct | gtccacgccc | ctggttctta  | gttccagccc  | cactcatagg | acactcatag | 4020 |
| cccaggaggg | crccgccrc  | aaccccaccc  | gctaaagtac  | ttggageggt | ctctccctcc | 4080 |
|            |            |             | ccaagagtgg  |            |            |      |
|            |            |             | ctccaacatg  |            |            |      |
|            |            |             | catcatggcc  |            |            |      |
|            |            |             | tgcggcgagc  |            |            |      |
| cggtaatacg | gttatccaca | gaatcagggg  | ataacgcagg  | aaagaacatg | tgagcaaaag | 4380 |
| gccagcaaaa | ggccaggaac | cgtaaaaagg  | ccgcgttgct  | ggcgtttttc | cataggetee | 4440 |
|            |            |             | gctcaagtca  |            |            |      |
| gactataaag | ataccaggcg | tttccccctg  | gaagctccct  | cgtgcgctct | cctgttccga | 4560 |
| cectgeeget | taccggatac | ctgtccgcct  | ttctcccttc  | gggaagcgtg | gcgctttctc | 4620 |
| atageteacg | ctgtaggtat | ctcagttcgg  | tgtaggtcgt  | tcgctccaag | ctgggctgtg | 4680 |
| tgcacgaacc | ccccgttcag | cccgaccgct  | gcgccttatc  | cggtaactat | cgtcttgagt | 4740 |
|            |            |             | tggcagcagc  |            |            |      |
| gagcgaggta | tgtaggcggt | gctacagagt  | tcttgaagtg  | gtggcctaac | tacggctaca | 4860 |
|            |            |             | tgctgaagcc  |            |            |      |
| ttggtagctc | ttgatccggc | aaacaaacca  | ccgctggtag  | cggtggtttt | tttgtttgca | 4980 |
|            |            |             | ctcaagaaga  |            |            |      |
| ggtctgacgc | tcagtggaac | gaaaactcac  | gttaagggat  | tttggtcatg | agattatcaa | 5100 |
| aaaggatctt | cacctagatc | cttttaaatt  | aaaaatgaag  | ttttaaatca | atctaaagta | 5160 |
| tatatgagta | aacttggtct | gacagttacc  | aatgcttaat  | cagtgaggca | cctatctcag | 5220 |
|            |            |             | cctgactcgg  |            |            |      |
| tgcctcgtga | agaaggtgtt | gctgactcat  | accaggcctg  | aatcgcccca | tcatccagcc | 5340 |
| agaaagtgag | ggagccacgg | ttgatgagag. | ctttgttgta  | ggtggaccag | ttggtgattt | 5400 |
| tgaacttttg | ctttgccacg | gaacggtctg  | cgttgtcggg  | aagatgcgtg | atctgatcct | 5460 |
| tcaactcagc | aaaagttcga | tttattcaac  | aaagccgccg  | tcccgtcaag | tcagcgtaat | 5520 |
| gctctgccag | tgttacaacc | aattaaccaa  | ttctgattag, | aaaaactcat | cgagcatcaa | 5580 |
| atgaaactgc | aatttattca | tatcaggatt  | atcaatacca  | tatttttgaa | aaagccgttt | 5640 |
| ctgtaatgaa | ggagaaaact | caccgaggca  | gttccatagg  | atggcaagat | cctggtatcg | 5700 |
| gtctgcgatt | ccgactcgtc | caacatcaat  | acaacctatt  | aatttcccct | cgtcaaaaat | 5760 |
| aaggttatca | agtgagaaat | caccatgagt  | gacgactgaa  | tccggtgaga | atggcaaaag | 5820 |
| cttatgcatt | tctttccaga | cttgttcaac  | aggccagcca  | ttacgctcgt | catcaaaatc | 5880 |
| actcgcatca | accaaaccgt | tattcattcg  | tgattgcgcc  | tgagcgagac | gaaatacgcg | 5940 |
| atcgctgtta | aaaggacaat | tacaaacagg  | aatcgaatgc  | aaccggcgca | ggaacactgc | 6000 |
| cagcgcatca | acaatatttt | cacctgaatc  | aggatattct  | tctaatacct | ggaatgctgt | 6060 |
| tttcccgggg | atcgcagtgg | tgagtaacca  | tgcatcatca  | ggagtacgga | taaaatgctt | 6120 |
| gatggtcgga | agaggcataa | attccgtcag  | ccagtttagt  | ctgaccatct | catctgtaac | 6180 |
| atcattggca | acgctacctt | tgccatgttt  | cagaaacaac  | tctggcgcat | cgggcttccc | 6240 |
| atacaatcga | tagattgtcg | cacctgattg  | cccgacatta  | tcgcgagccc | atttataccc | 6300 |
| atataaatca | gcatccatgt | tggaatttaa  | tegeggeete  | gagcaagacg | tttcccgttg | 6360 |
|            |            |             | gtttatgtaa  |            |            |      |
| tgatgatata | tttttatctt | gtgcaatgta  | acatcagaga  | ttttgagaca | caacgtggct | 6480 |
|            |            |             |             |            |            |      |

ttccccccc ccccattatt gaagcatta tcagggttat tgtctcatga gcggatacat 6540 atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 6600 gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat 6660 cacgaggccc tttcgtc 6677

<210> 122
<211> 6608
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 122

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaaccgte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tottatgcat gotatactgt tittggcttg gggcctatac acccccgctt cottatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagaetgae aeggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcggggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ccatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatge ecaceaceae cagtgtgeeg cacaaggeeg tggeggtagg gtatgtgtet 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctqccqcq 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980 gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040 gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100 acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160 cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220 tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agetgacece cetgtgegtg accetgeact geageaaceg caccategae tacaacaace 2340

gcaccgacaa catgggcggc gagatcaaga actgcagctt ccgatcgatc aactgcaaca 2400 ccagegecat caeccaggee tgececaagg tgagettega ceccatecee atecaetaet 2460 gegeeeege eggetacgee ateetgaagt geaacaacaa gaeetteaac ggeaeeggee 2520 cctgcaacaa cgtgagcacc atccagtgca cccacggcac caagcccgtg gtgagcaccc 2580 agctgctgct gaacggcagc ctggccgagg aggagatcat catccgcagc aagaacctga 2640 ccgacaacgt gaagaccatc atcgtgcacc tgaacgagag cgtggagatc aactgcaccc 2700 qccqatcqqc ccactqcaac atcagccgca ccgcctggaa caagaccctg caggaggtgg 2760 gcaagaaget ggccgagcac ttccccaaca aggccatcaa gttcgccaag cacagcggcg 2820 qcqacctqqa qatcaccacc cacagcttca actgccgcgg cgagttcttc tactgcaaca 2880 cocqatcqat ccctqccqc atcaaqcaqa tcatcaacat qtqqcaqqqc gtgggccgcq 2940 ccatgtacgc ccccccatc gagggcatcc tgacctgccg cagcaacatc accggcctgc 3000 tgctgacccg cgacggcggc accggcatgc acgacaccga gatcttccgc cccgagggcg 3060 gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg gagatcaagc 3120 ccctgggcat cgccccacc aaggccaagc ttaccgtgca ggtgcgccag ctgctgagcg 3180 gcatcgtgca gcagcagagc aacctgctgc gcgccatcga ggcccagcag cacatgctgc 3240 agetgacegt gtggggcate aageagetge agaceegegt getggeeate gagegetace 3300 tgcgcgacca gcagctcgag gacatctgga acaacatgac ctggatgcag tgggaccgcg 3360 agatcagcaa ctacaccaac accatctacc gcctgctgga ggacagccag aaccagcagg 3420 agaagaacga gcaggacctg ctggccctgg acaagtggca gaacctgtgg acctggttcg 3480 gcatcaccaa ctggctgtgg tacatcaaga tcttcatcaa gatcgtgggc ggcctgatcg 3540 gcctgcgcat catcttcgcc gtgctgagca tctgaacacg tgggatccag atctgctgtg 3600 cottotagtt gocagocate tgttgtttgc coetcoecg tgccttcctt gaccetggaa 3660 ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 3720 aggtgtcatt ctattctggg gggtggggtg gggcaggaca gcaaggggga ggattgggaa 3780 gacaatagca ggcatgctgg ggatgcggtg ggctctatgg gtacccaggt gctgaagaat 3840 tgacccggtt cetectggge cagaaagaag caggeacate ceettetetg tgacacacec 3900 tgtccacgcc cctggttctt agttccagcc ccactcatag gacactcata gctcaggagg 3960 qctccqcctt caatcccacc cgctaaagta cttggagegg tctctccctc cctcatcagc 4020 ccaccaaacc aaacctagcc tccaagagtg ggaagaaatt aaagcaagat aggctattaa 4080 gtgcagaggg agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat 4140 tttaaggcca tgatttaagg ccatcatggc cttaatcttc cgcttcctcg ctcactgact 4200 cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac 4260 ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 4320 aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg 4380 acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa 4440 gataccagge gtttececet ggaageteee tegtgegete teetgtteeg accetgeege 4500 ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac 4560 gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 4620 cccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg 4680 taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 4740 atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa 4800 cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaaga gttggtagct 4860 cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga 4920 ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg 4980 ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct 5040 tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 5100 aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc 5160 aagaaggtgt tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga 5280 gggagccacg gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt 5340 gctttgccac ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag 5400 caaaaqttcq atttattcaa caaagccgcc qtcccqtcaa gtcagcgtaa tgctctgcca 5460 gtgttacaac caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg 5520 caatttattc atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga 5580 aggagaaaac tcaccgaggc agttccatag gatggcaaga tcctggtatc ggtctgcgat 5640

```
tccqactcqt ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc 5700
aagtgagaaa tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat 5760
ttotttocag acttgttoaa caggocagoo attacgotog toatcaaaat cactogoato 5820
aaccaaaccg ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt 5880
aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc 5940
aacaatattt tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg 6000
gatcqcaqtq gtqaqtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg 6060
aaqaqqcata aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc 6120
aacgctacct ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg 6180
atagattgtc gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc 6240
agcatccatg ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct 6300
cataacaccc cttqtattac tqtttatgta agcagacagt tttattgttc atgatgatat 6360
atttttatct tqtqcaatqt aacatcagag attttgagac acaacgtggc tttcccccc 6420
cccccattat tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg 6480
tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga 6540
cqtctaaqaa accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc 6600
                                                                  6608
ctttcgtc
```

<210> 123 <211> 6689 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 123 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 eccgectgge tgacegeeca acgaeeceeg eccattgaeg teaataatga egtatgttee 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tottatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaeteece 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 agggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeac 1560

|   |            |            |            |            | tggcggtagg |            |        |
|---|------------|------------|------------|------------|------------|------------|--------|
|   |            |            |            |            | atggaagact |            |        |
|   |            |            |            |            | aagagtcaga |            |        |
|   |            |            |            |            | cagtactcgt |            |        |
|   |            |            |            |            | tcctttccat |            |        |
|   | tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatcgcca | 1920   |
|   | ccatgcgcgt | ggagggcatc | cagcgcaact | ggaagcagtg | gtggatctgg | ggcatcctgg | 1980   |
|   | gcttctggat | ggtgatgatc | tacaacgtgc | gcggcaacct | gtgggtgacc | gtgtactacg | 2040   |
|   | gcgtgcccgt | gtggaaggag | gccaagacca | ccctgttctg | cgccagcgac | gccaaggcct | 2100 . |
|   | acgacgccga | ggtgcacaac | gtgtgggcca | cccacgcctg | cgtgcccacc | gaccccaacc | 2160   |
|   | cccaggagat | ggtgctggag | aacgtgaccg | agaacttcaa | catgtgggag | aacgacatgg | 2220   |
|   | tggagcagat | gcaccaggac | atcatcagcc | tgtgggacca | gagcctgaag | ccctgcgtga | 2280   |
|   | agctgacccc | cctgtgcgtg | accctgcact | gcagcaaccg | caccatcgac | tacaacaacc | 2340   |
|   | gcaccgacaa | catgggcggc | gagatcaaga | actgcagctt | ccgatcgatc | aactgcaaca | 2400   |
|   | ccagcgccat | cacccaggcc | tgccccaagg | tgagcttcga | ccccatcccc | atccactact | 2460   |
|   | gcgcccccgc | cggctacgcc | atcctgaagt | gcaacaacaa | gaccttcaac | ggcaccggcc | 2520   |
|   | cctgcaacaa | cgtgagcacc | atccagtgca | cccacggcac | caagcccgtg | gtgagcaccc | 2580   |
|   | agctgctgct | gaacggcagc | ctggccgagg | aggagatcat | catccgcagc | aagaacctga | 2640   |
|   | ccgacaacgt | gaagaccatc | atcgtgcacc | tgaacgagag | cgtggagatc | aactgcaccc | 2700   |
|   |            |            |            |            | cggccaggcc |            |        |
|   |            |            |            |            | catcagccgc |            |        |
|   |            |            |            |            | cttccccaac |            |        |
|   |            |            |            |            | ccacagette |            |        |
|   |            |            |            |            | catcaagcag |            |        |
|   |            |            |            |            | cgagggcatc |            |        |
|   |            |            |            |            | caccggcatg |            |        |
|   |            |            |            |            | gcgcagcgag |            |        |
|   |            |            |            |            | caaggccaag |            |        |
|   |            |            |            |            | caacctgctg |            |        |
|   |            |            |            |            | caagcagctg |            |        |
|   |            |            |            |            | ggacatctgg |            |        |
|   |            |            |            |            | caccatctac |            |        |
|   |            |            |            |            | gctggccctg |            |        |
|   |            |            |            |            | gtacatcaag |            |        |
|   |            |            |            |            | cgtgctgagc |            |        |
|   |            |            |            |            | ctgttgtttg |            |        |
|   |            |            |            |            | tttcctaata |            |        |
|   |            |            |            |            | ggggtggggt |            |        |
|   |            |            |            |            | gggatgcggt |            |        |
|   |            |            |            |            | ccagaaagaa |            |        |
|   |            |            |            |            | tagttccagc |            |        |
|   | ggacactcat | agctcaggag | ggctccgcct | tcaatcccac | ccgctaaagt | acttggagcg | 4080   |
|   |            |            |            |            | ctccaagagt |            |        |
| • |            |            |            |            | gcctccaaca |            |        |
|   |            |            |            |            | gccatcatgg |            |        |
|   |            |            |            |            | gctgcggcga |            |        |
|   | ctcactcaaa | ggcggtaata | cggttatcca | cagaatcagg | ggataacgca | ggaaagaaca | 4380   |
|   | tgtgagcaaa | aggccagcaa | aaggccagga | accgtaaaaa | ggccgcgttg | ctggcgtttt | 4440   |
|   |            |            |            |            | acgctcaagt |            |        |
|   |            |            |            |            | tggaagetee |            |        |
|   |            |            |            |            | ctttctccct |            |        |
|   |            |            |            |            | ggtgtaggtc |            |        |
|   |            |            |            |            | ctgcgcctta |            |        |
|   |            |            |            |            | actggcagca |            |        |
|   |            |            |            |            | gttcttgaag |            |        |
|   |            |            |            |            | _          |            |        |

```
actacgqcta cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct 4920
teggaaaaag agttggtage tettgateeg geaaacaaac cacegetggt ageggtggtt 4980
tttttqtttq caaqcaqcag attacqcqca qaaaaaaagg atctcaagaa gatcctttga 5040
tetittetae ggggtetgae geteagtgga acgaaaacte acgttaaggg attittggtca 5100
tqaqattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat 5160
caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg 5220
qqcqctqaqq tctgcctcgt gaagaaggtg ttgctgactc ataccaggcc tgaatcgccc 5340
catcatccag ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc 5400
agttqqtqat tttqaacttt tqctttqcca cqqaacqqtc tqcqttqtcq gqaagatqcq 5460
tqatctqatc cttcaactca qcaaaaqttc qatttattca acaaagccgc cgtcccgtca 5520
agtcagcgta atgctctgcc agtgttacaa ccaattaacc aattctgatt agaaaaactc 5580
atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg 5640
aaaaaqccqt ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag 5700
atcctggtat cggtctgcga ttccgactcg tccaacatca atacaaccta ttaatttccc 5760
ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga 5820
gaatggcaaa agcttatgca tttctttcca gacttgttca acaggccagc cattacgctc 5880
qtcatcaaaa tcactcgcat caaccaaacc gttattcatt cgtgattgcg cctgagcgag 5940
acgaaatacg cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaacoggcg 6000
caggaacact gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac 6060
ctggaatgct gttttcccgg ggatcgcagt ggtgagtaac catgcatcat caggagtacg 6120
gataaaatgc ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat 6180
ctcatctgta acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc 6240
ategggette ccatacaate gatagattgt egeacetgat tgeeegacat tategegage 6300
ccatttatac ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga 6360
cgtttcccgt tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag 6420
ttttattqtt catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga 6480
cacaacgtgg ctttccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 6540
qaqqqatac atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt 6600
tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa 6660
aaataggcgt atcacgaggc cctttcgtc
```

```
<210> 124
<211> 6785
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVRl012x/s containing HIV genes
<400> 124
```

tegegegtt eggtgatgac ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 eagettgtet gtaageggat geeggagea gacaageeeg teagggegeg teagegggtg 120 teggegggtg teggggetgg ettaactatg eggeateaga geagattgta etgagagtge 180 accatatgeg gtgtgaaata eegecacagat gegtaaggag aaaataceege ateagattgg 240 etattggeea tegeateegt tgateeata teataatatg tacattata teggeteatg 300 tecaacatta eegecatgtt gacattgatt attgactagt tattaatagt aateaattac 360 ggggteatta gteatagee eataatgga gtteegegtt acataactta eggtaaatgg 420 eeegeetgge tgacegeea acgaeeeegg eeeattgge teaatagga ettteeattg aegteaatgg gtggagtatt taeggtaaac 540 tegeeggtaaa tegeeegeet ggeattatge eeagtacatg aeettatggg aeetteetae 660 teggeagtae atetaegtat tagteatege tattaeeatg gtggatggt tettggeagta 720

WO 02/32943 PCT/US01/25721 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780

|            | cgtggatagc |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | agtttgtttt |            |            |            |            |      |
|            | ttgacgcaaa |            |            |            |            |      |
|            | gtgaaccgtc |            |            |            |            |      |
|            | cgggaccgat |            |            |            |            |      |
|            | aagagtgacg |            |            |            |            |      |
| tcttatgcat | gctatactgt | ttttggcttg | gggcctatac | acccccgctt | ccttatgcta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg | tgggttattg | accattattg | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260 |
| tattggctat | atgccaatac | tctgtccttc | agagactgac | acggactctg | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacgccgt | ccccgtgcc  | 1380 |
| cgcagttttt | attaaacata | gcgtgggatc | tccacgcgaa | tctcgggtac | gtgttccgga | 1440 |
| catgggctct | tctccggtag | cggcggagct | tccacatccg | agccctggtc | ccatgcctcc | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620 |
|            | gtggagattg |            |            |            |            |      |
|            | atgcaggcag |            |            |            |            |      |
|            | tgttaacggt |            |            |            |            |      |
|            | gacataatag |            |            |            |            |      |
|            | gtcgtcgaca |            |            |            |            |      |
|            | ggagggcatc |            |            |            |            |      |
|            | ggtgatgatc |            |            |            |            |      |
|            | gtggaaggag |            |            |            |            |      |
|            | ggtgcacaac |            |            |            |            |      |
|            | ggtgctggag |            |            |            |            |      |
|            | gcaccaggac |            |            |            |            |      |
|            | cctgtgcgtg |            |            |            |            |      |
|            | catgggcggc |            |            |            |            |      |
|            | cgagaaggtg |            |            |            |            |      |
|            | caacaccagc |            |            |            |            |      |
|            | ccccaaggtg |            |            |            |            |      |
|            | cctgaagtgc |            |            |            |            |      |
|            | ccagtgcacc |            |            |            |            |      |
|            | ggccgaggag |            |            |            |            |      |
|            | cgtgcacctg |            |            |            |            |      |
|            | cagccgcacc |            |            |            |            |      |
|            | ccccaacaag |            |            |            |            |      |
|            | cagcttcaac |            |            |            |            |      |
|            | ctacaccccc |            |            |            |            |      |
|            | ctgccgcatc |            |            |            |            |      |
|            | ccccatcgag |            |            |            |            |      |
|            | cggcggcacc |            |            |            |            |      |
|            | caactggcgc |            |            |            |            |      |
|            | ccccaccaag |            |            |            |            |      |
|            | gcagagcaac |            |            |            |            |      |
|            | gggcatcaag |            |            |            |            |      |
|            | gctcgaggac |            |            |            |            |      |
|            | caccaacacc |            |            |            |            |      |
|            | ggacctgctg |            |            |            |            |      |
|            | gctgtggtac |            |            |            |            |      |
|            | cttcgccgtg |            |            |            |            |      |
|            | agccatctgt |            |            |            |            |      |
|            | ctgtcctttc |            |            |            |            |      |
|            | ttctgggggg |            |            |            |            |      |
|            |            |            |            |            |            |      |
| aacaycayyc | atgctgggga | -3-33-3336 | cccacyyyca | cccaygract | yaayaattyd | 4020 |

```
cccggttcct cctgggccag aaagaagcag gcacatcccc ttctctgtga cacaccctgt 4080
ccacgccct ggttcttagt tccagccca ctcataggac actcatagct caggagggct 4140
cegeetteaa teecaceege taaagtaett ggageggtet eteecteeet cateageeca 4200
ccaaaccaaa cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagtg 4260
cagagggaga gaaaatgcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt 4320
aaggccatga tttaaggcca tcatggcctt aatcttccgc ttcctcgctc actgactcgc 4380
tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt 4440
tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg 4500
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg 4560
agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 4620
accaggegtt tececetgga ageteceteg tgegetetee tgtteegace etgeegetta 4680
eeggataeet gteegeettt eteeettegg gaagegtgge gettteteat ageteaeget 4740
gtaggtatet cagtteggtg taggtegtte getecaaget gggetgtgtg caegaacee 4800
cogtteagec egacegotge geettateeg gtaactateg tottgagtec aacceggtaa 4860
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 4920
taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag 4980
tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 5040
gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 5100
cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 5160
agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 5220
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 5280
cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 5340
ttcgttcatc catagttgcc tgactcgggg ggggggggc ctgaggtctg cctcgtgaag 5400
aaggtgttgc tgactcatac caggcctgaa tcgccccatc atccagccag aaagtgaggg 5460
agccacggtt gatgagagct ttgttgtagg tggaccagtt ggtgattttg aacttttgct 5520
ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat ctgatccttc aactcagcaa 5580
aagttegatt tatteaacaa ageegeegte eegtcaagte agegtaatge tetgeeagtg 5640
ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa 5700
tttattcata tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg 5760
agaaaactca ccgaggcagt tccataggat ggcaagatcc tggtatcggt ctgcgattcc 5820
gactegteca acateaatac aacetattaa titteeeeteg teaaaaataa ggttateaag 5880
tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc 5940
tttccaqact tgttcaacag gccaqccatt acgctcgtca tcaaaatcac tcgcatcaac 6000
caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 6060
aggacaatta caaacaggaa togaatgcaa coggogoagg aacactgcca gogoatcaac 6120
aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 6180
cgcagtggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6240
aggcataaat teegteagee agtttagtet gaccatetea tetgtaacat cattggcaac 6300
gctacctttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6360
gattqtcqca cctqattqcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6420
atccatgttg gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6480
aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt 6540
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6600
ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6660
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6720
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6780
                                                                  6785
tcgtc
```

<sup>&</sup>lt;210> 125

<sup>&</sup>lt;211> 6716

<sup>&</sup>lt;212> DNA

<sup>&</sup>lt;213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> .125 tegegegttt eggtgatgae ggtgaaaace tetgacacat qeaqeteecq qaqacqqtca 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctqqc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtacege etatagaete tataggeaca eccetttgge 1080 tettatgcat getatactgt ttttggettg gggeetatac acceegett cettatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggactetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980 gettetggat ggtgatgate tacaacgtge geggeaacet gtgggtgace gtgtactacq 2040 gegtgeeegt gtggaaggag geeaagacea ceetgttetg egeeagegae geeaaggeet 2100 acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160 cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220 tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agetgacece cetgtgegtg accetgeact geageaaceg caccategae tacaacaace 2340 gcaccgacaa catgggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400 gegacaageg egagaaggtg caegecetgt tetacegeet ggacategtg cecetgaaga 2460 acgagagcag caacaccagc ggcgactacc gcctgatcaa ctgcaacacc agcgccatca 2520 cccaggcctg ccccaaggtg agcttcgacc ccatccccat ccactactgc gcccccgccg 2580 gctacgccat cctgaagtgc aacaacaaga ccttcaacgg caccggcccc tgcaacaacg 2640 tgagcaccat ccagtgcacc cacggcacca agcccgtggt gagcacccag ctgctgctga 2700 acggcagcct ggccgaggag gagatcatca tccgcagcaa gaacctgacc gacaacgtga 2760 agaccatcat cgtgcacctg aacgagagcg tggagatcaa ctgcacccgc cgatcggccc 2820 actgcaacat cagccgcacc gcctggaaca agaccctgca ggaggtgggc aagaagctgg 2880 ccgagcactt ccccaacaag gccatcaagt tcgccaagca cagcggcggc gacctggaga 2940 tcaccaccca cagcttcaac tgccgcggcg agttcttcta ctgcaacacc cgatcgatcc 3000 cetgeegeat caageagate ateaacatgt ggcagggegt gggcegegee atgtacgeec 3060

| cccccatcga | gggcatcctg | acctgccgca  | gcaacatcac | cggcctgctg | ctgacccgcg | 3120 |
|------------|------------|-------------|------------|------------|------------|------|
| acggcggcac | cggcatgcac | gacaccgaga  | tetteegeee | cgagggcggc | gacatgcgcg | 3180 |
| acaactggcg | cagcgagctg | tacaagtaca  | aggtggtģga | gatcaagccc | ctgggcatcg | 3240 |
|            |            |             | tgcgccagct |            |            |      |
|            |            |             | cccagcagca |            |            |      |
|            |            |             | tggccatcga |            |            |      |
| agctcgagga | catctggaac | aacatgacct  | ggatgcagtg | ggaccgcgag | atcagcaact | 3480 |
| acaccaacac | catctaccgc | ctgctggagg  | acagccagaa | ccagcaggag | aagaacgagc | 3540 |
|            |            |             | acctgtggac |            |            |      |
| ggctgtggta | catcaagatc | ttcatcaaga  | tcgtgggcgg | cctgatcggc | ctgcgcatca | 3660 |
| tcttcgccgt | gctgagcatc | tgaacacgtg  | ggatccagat | ctgctgtgcc | ttctagttgc | 3720 |
| cagccatctg | ttgtttgccc | ctccccgtg   | ccttccttga | ccctggaagg | tgccactccc | 3780 |
| actgtccttt | cctaataaaa | tgaggaaatt  | gcatcgcatt | gtctgagtag | gtgtcattct | 3840 |
| attctggggg | gtggggtggg | gcaggacagc  | aagggggagg | attgggaaga | caatagcagg | 3900 |
| catgctgggg | atgcggtggg | ctctatgggt  | acccaggtgc | tgaagaattg | acceggttee | 3960 |
| tcctgggcca | gaaagaagca | ggcacatccc  | cttctctgtg | acacaccctg | tccacgcccc | 4020 |
| tggttcttag | ttccagcccc | actcatagga  | cactcatage | tcaggagggc | tccgccttca | 4080 |
| atcccacccg | ctaaagtact | tggagcggtc  | tetecetece | tcatcagccc | accaaaccaa | 4140 |
| acctagcctc | caagagtggg | aagaaattaa  | agcaagatag | gctattaagt | gcagagggag | 4200 |
|            |            |             | tgagagaaat |            |            |      |
| atttaaggcc | atcatggcct | taatcttccg  | cttcctcgct | cactgactcg | ctgcgctcgg | 4320 |
| tcgttcggct | gcggcgagcg | gtatcagctc  | actcaaaggc | ggtaatacgg | ttatccacag | 4380 |
|            |            |             | gagcaaaagg |            |            |      |
| gtaaaaaggc | cgcgttgctg | gcgtttttcc  | ataggctccg | ccccctgac  | gagcatcaca | 4500 |
| aaaatcgacg | ctcaagtcag | aggtggcgaa  | acccgacagg | actataaaga | taccaggcgt | 4560 |
|            |            |             | ctgttccgac |            |            |      |
| tgtccgcctt | tctcccttcg | ggaagcgtgg  | cgctttctca | tageteacge | tqtaqqtatc | 4680 |
| tcagttcggt | gtaggtcgtt | cgctccaagc  | tgggctgtgt | gcacqaaccc | cccqttcaqc | 4740 |
|            |            |             | gtcttgagtc |            |            |      |
|            |            |             | ggattagcag |            |            |      |
| ctacagagtt | cttgaagtgg | tggcctaact  | acggctacac | tagaagaaca | gtatttggta | 4920 |
| tctgcgctct | gctgaagcca | gttaccttcg  | gaaaaagagt | tggtagctct | tgatccggca | 4980 |
| aacaaaccac | cgctggtagc | ggtggttttt  | ttgtttgcaa | gcagcagatt | acqcqcaqaa | 5040 |
| aaaaaggatc | tcaagaagat | cctttgatct  | tttctacggg | gtctgacgct | cagtggaacg | 5100 |
|            |            |             | gattatcaaa |            |            |      |
|            |            |             | tctaaagtat |            |            |      |
|            |            |             | ctatctcagc |            |            |      |
| ccatagttgc | ctgactcggg | gggggggggc  | gctgaggtct | qcctcqtqaa | gaaggtgttg | 5340 |
| ctgactcata | ccaggcctga | atcgccccat  | catccagcca | gaaagtgagg | qaqccacqqt | 5400 |
| tgatgagagc | tttgttgtåg | gtggaccagt  | tggtgatttt | gaacttttgc | tttqccacqq | 5460 |
| aacggtctgc | gttgtcggga | agatgcgtga  | tctgatcctt | caactcagca | aaagttcgat | 5520 |
|            |            |             | cagcgtaatg |            |            |      |
|            |            |             | gagcatcaaa |            |            |      |
| atcaggatta | tcaataccat | atttttqaaa  | aagccgtttc | totaatoaao | gagaaaactc | 5700 |
| accgaggcag | ttccatagga | tggcaaqatc  | ctggtatcgg | tctqcqattc | cgactcgtcc | 5760 |
| aacatcaata | caacctatta | atttcccctc  | gtcaaaaata | aggttatcaa | gtgagaaatc | 5820 |
| accatgagtg | acqactqaat | ccqqtqaqaa  | tggcaaaagc | ttatgcattt | ctttccagac | 5880 |
| ttgttcaaca | ggccagccat | tacqctcqtc  | atcaaaatca | ctcgcatcaa | ccaaaccgtt | 5940 |
| attcattcqt | gattgcgcct | qaqcqaqacq  | aaatacgcga | tcoctottaa | aaggacaatt | 6000 |
| acaaacagga | atcqaatqca | acconcocad  | gaacactgcc | agcgcatcaa | caatattttc | 6060 |
| acctgaatca | ggatattctt | ctaataccto  | gaatgctgtt | ttcccaaaaa | tcqcaataat | 6120 |
| gagtaaccat | gcatcatcag | gagtacqqat  | aaaatgcttg | atggtcggaa | gaggcataaa | 6180 |
| ttccgtcagc | cagtttagtc | tgaccatete  | atctgtaaca | tcattoggaa | cactacetth | 6240 |
| qccatqtttc | agaaacaact | ctggcgcate  | gggcttccca | tacaarcoar | agattotcoc | 6300 |
| acctgattgc | ccgacattat | cqcqaqccca  | tttataccca | tataaatcag | catecators | 6360 |
|            | J          | J J. J. 224 |            |            |            | 3330 |

```
ggaatttaat cgcggcctcg agcaagacgt ttcccgttga atatggctca taacacccct 6420 tgtattactg tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg 6480 tgcaatgtaa catcagagat tttgagacac aacgtggctt tcccccccc cccattattg 6540 aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 6600 taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac 6660 cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 6716
```

<210> 126 <211> 6797 <212> DNA <213> Artificial Sequence <223> plasmid pVR1012x/s containing HIV genes <400> 126 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgog gtgtgaaata cogcacagat gogtaaggag aaaatacogc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaaqtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tggqcggtag gcqtqtacqq tggqagqtct atataaqcaq 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcqqgtac qtqttccqqa 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860

tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980 gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040 gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100 acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgccacc gaccccaacc 2160 cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220

tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280 agetgacece cetgtgegtg accetgeact geageaaceg caccategae tacaacaace 2340 gcaccgacaa catgggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400 gcqacaaqcq cgaqaaggtg cacgccctgt tctaccgcct ggacatcgtg cccctgaaga 2460 acgagagcag caacaccagc ggcgactacc gcctgatcaa ctgcaacacc agcgccatca 2520 eccaggeetg ceccaaggtg agettegace ceatececat ceactactge gecceegeeg 2580 gctacgccat cctgaagtgc aacaacaaga ccttcaacgg caccggcccc tgcaacaacg 2640 tgagcaccat ccagtgcacc cacggcacca agcccgtggt gagcacccag ctgctgctga 2700 acggcageet ggeegaggag gagateatea teegeageaa gaacetgaee gacaaegtga 2760 agaccatcat cgtgcacctg aacgagagcg tggagatcaa ctgcacccgc cccaacaaca 2820 acaccegcaa gagcateege ateggeeeeg geeaggeett etaegeeace ggegagatea 2880 teggegacat cegecaggee caetgeaaca teageegeae egeetggaae aagaeeetge 2940 aggaggtggg caagaagctg geegageact teeceaacaa ggeeateaag ttegeeaage 3000 acageggegg egaeetggag ateaecacee acagetteaa etgeegegge gagttettet 3060 actgcaacac ccgatcgatc ccctgccgca tcaagcagat catcaacatg tggcagggcg 3120 tgggccgcgc catgtacgcc cccccatcg agggcatcct gacctgccgc agcaacatca 3180 ceggeetget getgaeeege gaeggeggea eeggeatgea egaeaeegag atetteegee 3240 ccgagggcgg cgacatgcgc gacaactggc gcagcgagct gtacaagtac aaggtggtgg 3300 agatcaagee cetgggcate geececacca aggccaaget tacegtgcag gtgegccage 3360 tgctgagcgg catcgtgcag cagcagagca acctgctgcg cgccatcgag gcccagcagc 3420 acatgctgca gctgaccgtg tggggcatca agcagctgca gacccgcgtg ctggccatcg 3480 agogotacot gogogacoag cagotogagg acatotggaa caacatgaco tggatgcagt 3540 gggaccgcga gatcagcaac tacaccaaca ccatctaccg cctgctggag gacagccaga 3600 accagcagga gaagaacgag caggacctgc tggccctgga caagtggcag aacctgtgga 3660 cctggttcgg catcaccaac tggctgtggt acatcaagat cttcatcaag atcgtgggcg 3720 gcctgategg cctgcgcate atcttegccg tgctgagcat ctgaacacgt gggatecaga 3780 tetgetgtge ettetagttg ceagecatet gttgtttgce cetececegt geetteettg 3840 accetggaag gtgccactce cactgteett tectaataaa atgaggaaat tgcategcat 3900 tqtctqaqta qqtqtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag 3960 gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatggg tacccaggtg 4020 ctgaagaatt gacccggttc ctcctgggcc agaaagaagc aggcacatcc ccttctctgt 4080 gacacaccct gtccacgccc ctggttctta gttccagccc cactcatagg acactcatag 4140 ctcaggaggg ctccgccttc aatcccaccc gctaaagtac ttggagcggt ctctccctcc 4200 ctcatcagcc caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata 4260 ggctattaag tgcagaggga gagaaaatgc ctccaacatg tgaggaagta atgagagaaa 4320 tcatagaatt ttaaggccat gatttaaggc catcatggcc ttaatcttcc gcttcctcgc 4380 teactgacte getgegeteg gtegttegge tgeggegage ggtateaget cacteaaagg 4440 cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 4500 gccagcaaaa ggccaggaac cgtaaaaaagg ccgcgttgct ggcgtttttc cataggctcc 4560 qccccctqa cqaqcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 4620 gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 4680 ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc 4740 atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg 4800 tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 4860 ccaacceggt aagacacgac ttategceac tggcagcagc cactggtaac aggattagca 4920 gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 4980 chagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 5040 ttggtagete ttgateegge aaacaaacca eegetggtag eggtggtttt tttgtttgca 5100 agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 5160 ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 5220 aaaggatett eacetagate etittaaatt aaaaatgaag tittaaatea atetaaagta 5280 tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag 5340 tgcctcgtga agaaggtgtt gctgactcat accaggcctg aatcgcccca tcatccagcc 5460 agaaagtgag ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt 5520

```
tgaacttttg ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatect 5580
teaacteage aaaagttega tttatteaac aaageegeeg teeegteaag teagegtaat 5640
gctctgccag tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa 5700
atgaaactgc aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt 5760
ctgtaatgaa ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg 5820
gtctgcgatt ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat 5880
aaggttatca agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag 5940
cttatgcatt tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc 6000
actogoatca accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg 6060
atcqctqtta aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc 6120
cagogoatca acaatatttt cacotgaatc aggatattot totaatacot ggaatgotgt 6180
tttcccgggg atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaatgctt 6240
gatggtegga agaggeataa atteegteag eeagtttagt etgaceatet catetgtaac 6300
atcattggca acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc 6360
atacaatcga tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc 6420
atataaatca gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg 6480
aatatggctc ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca 6540
tgatgatata tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct 6600
ttccccccc ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat 6660
atttgaatgt atttagaaaa ataaacaaat aggggtteeg egcacattte eeegaaaagt 6720
gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat 6780
cacgaggccc tttcgtc
```

<210> 127 <211> 6718 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tegegegtht eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattqqcca ttqcatacqt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta ceqecatqtt qacattqatt attqactaqt tattaataqt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccqcctqqc tqaccqccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 checquecca thqacqcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaaccgtc agategectg gagacgccat ccacgetgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tettatgeat getatactgt tittiggettg gggeetatac acceegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260

| WO 02/3294 | 3          |            |            |            | PCT/US      | 801/257 |
|------------|------------|------------|------------|------------|-------------|---------|
| tattggctat | atgccaatac | tctgtccttc | agagactgac | acggactctg | tatttttaca  | 1320    |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacgccgt | ccccqtqcc   | 1380    |
| cgcagttttt | attaaacata | gcgtgggatc | tccacgcgaa | tctcgggtac | gtgttccgga  | 1440    |
| catgggctct | tctccqqtaq | cqqcqqaqct | tccacatccg | agecetgate | ccatqcctcc  | 1500    |
|            |            |            | ctcctaacag |            |             |         |
|            |            |            | cacaaggccg |            |             |         |
|            |            |            | gctgacgcag |            |             |         |
|            |            |            | gtattctgat |            |             |         |
|            |            |            | gtagtctgag |            |             |         |
|            |            |            | aacagactgt |            |             |         |
|            |            |            | gatatcgcgg |            |             |         |
|            |            |            | ggccgcagtg |            |             |         |
|            |            |            | ttggtaacat |            |             |         |
|            |            |            | ccctgttctg |            |             |         |
|            |            |            | cccacgcttg |            |             |         |
|            |            |            | aaaacttcaa |            |             |         |
|            |            |            | tgtgggacca |            |             |         |
|            |            |            | gcaccaacgc |            |             |         |
|            |            |            | agcagggtta |            |             |         |
|            |            |            | actccaacaa |            |             |         |
|            |            |            | gcccgaaagt |            |             |         |
|            |            |            | tectgaaatg |            |             |         |
|            |            |            | ttcagtgcac |            |             |         |
|            |            |            | tggctgaaaa |            |             |         |
|            |            |            | tcgttcacct |            |             |         |
|            |            |            | aatccatgcg |            |             |         |
|            |            |            | tccgtcaggc |            |             |         |
|            |            |            | aagaaaaact |            |             |         |
|            |            |            | gtggtgacct |            |             |         |
|            |            |            | acaccacccg |            |             |         |
|            |            |            | gtatcaaaca |            |             |         |
| atattaatca | toctatotac | actccaccaa | tcgctggtaa | catcacctcc | aaatccaaca  | 3180    |
|            |            |            | gtgaagacaa |            |             |         |
|            |            |            | ggcgttccga |            |             |         |
|            |            |            | ccggtgctaa |            |             |         |
|            |            |            | ccaacctgct |            |             |         |
|            |            |            | tcaaacagct |            |             |         |
|            |            |            | agatetggaa |            |             |         |
|            |            |            | ccatctaccg |            |             |         |
|            |            |            | tggctctgga |            |             |         |
|            |            |            | gaggatccag |            |             |         |
|            |            |            | tgccttcctt |            |             |         |
|            |            |            | ttgcatcgca |            |             |         |
|            |            |            | gcaaggggga |            |             |         |
|            |            |            | gtacccaggt |            |             |         |
|            |            |            | cccttctctg |            |             |         |
|            |            |            | gacactcata |            |             |         |
|            |            |            | tetetecete |            |             |         |
|            |            |            | aaagcaagat |            |             |         |
|            |            |            | aatgagagaa |            |             |         |
|            |            |            | cgcttcctcg |            |             |         |
|            |            |            | tcactcaaag |            |             |         |
|            |            |            | gtgagcaaaa |            |             |         |
|            |            |            | ccataggctc |            |             |         |
|            |            |            | aaacccgaca |            |             |         |
| Jacana     | -500000900 | -3-33-33-3 |            | 2200000000 | 24 caccadac | 1500    |

```
gtttccccct ggaageteee tegtgegete teetgtteeg accetgeege ttaceggata 4620
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 4680
teteagting gigtagging thegeteeaa geigggeigt gigcaegaac ceeeegtica 4740
gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga 4800
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4860
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4920
tatctgcgct ctgctgaagc cagttacctt cqqaaaaaga gttggtagct cttgatccqq 4980
caaacaaacc accqctqqta qcqqtqqttt ttttqtttqc aaqcaqcaqa ttacqcqcaq 5040
aaaaaaagga teteaagaag ateetttgat ettttetaeg gggtetgaeg eteagtggaa 5100
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 5160
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5220
tgacagttac caatgettaa teagtgagge acetatetea gegatetgte tatttegtte 5280
atccatagtt gcctgactcg ggggggggg gcgctgaggt ctgcctcgtg aagaaggtgt 5340
tgctgactca taccaggect gaategeece atcatecage cagaaagtga gggagecacq 5400
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5460
ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag caaaagttcg 5520
atthattcaa caaageegee gteeegteaa gteagegtaa tgetetgeea gtgttacaac 5580
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttattc 5640
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5700
teacegagge agttecatag gatggeaaga teetggtate ggtetgegat teegactegt 5760
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5820
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5880
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 5940
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 6000
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 6060
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg gatcgcagtg 6120
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6180
aattoogtoa gooagtttag totgaccato toatotgtaa catoattggo aacgotacot 6240
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6300
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6360
ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6420
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6480
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc cccccattat 6540
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6600
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6660
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc
```

```
<210> 128
<211> 6586
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 128
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccggagca gacaagcccg tcagggggg tcagcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtg 120
ttggcgggt tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
```

|            |            | P          |             |             |            |      |
|------------|------------|------------|-------------|-------------|------------|------|
|            |            |            | cccattgacg  |             |            |      |
|            |            |            | acgtcaatgg  |             |            |      |
|            |            |            | tatgccaagt  |             |            |      |
|            |            |            | ccagtacatg  |             |            |      |
| ttggcagtac | atctacgtat | tagtcatcgc | tattaccatg  | gtgatgcggt  | tttggcagta | 720  |
| catcaatggg | cgtggatagc | ggtttgactc | acggggattt  | ccaagtctcc  | accccattga | 780  |
|            |            |            | tcaacgggac  |             |            |      |
| ctccgcccca | ttgacgcaaa | tgggcggtag | gcgtgtacgg  | tgggaggtct  | atataagcag | 900  |
| agctcgttta | gtgaaccgtc | agatcgcctg | gagacgccat  | ccacgctgtt  | ttgacctcca | 960  |
|            |            |            | cggccgggaa  |             |            |      |
| tccccgtgcc | aagagtgacg | taagtaccgc | ctatagactc  | tataggcaca  | cccctttggc | 1080 |
| tcttatgcat | gctatactgt | ttttggcttg | gggcctatac  | acccccgctt  | ccttatgcta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg | tgggttattg  | accattattg  | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc | cataacatgg  | ctctttgcca  | caactatctc | 1260 |
| tattggctat | atgccaatac | tctgtccttc | agagactgac  | acggactctg  | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca  | acaacgccgt  | cccccgtgcc | 1380 |
| cgcagttttt | attaaacata | gcgtgggatc | tccacgcgaa  | tctcgggtac  | gtgttccgga | 1440 |
| catgggctct | tctccggtag | cggcggagct | tccacatccg  | agccctggtc  | ccatgcctcc | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag  | tggaggccag  | acttaggcac | 1560 |
|            |            |            | cacaaggccg  |             |            |      |
|            |            |            | gctgacgcag  |             |            |      |
|            |            |            | gtattctgat  |             |            |      |
|            |            |            | gtagtctgag  |             |            |      |
|            |            |            | aacagactgt  |             |            |      |
|            |            |            | gatatcgcgg  |             |            |      |
|            |            |            | ggccgcagtg  |             |            |      |
|            |            |            | ttggtaacat  |             |            |      |
|            |            |            | ccctgttctg  |             |            |      |
|            |            |            | cccacgcttg  |             |            |      |
|            |            |            | aaaacttcaa  |             |            |      |
|            |            |            | tgtgggacca  |             |            |      |
|            |            |            | gcaccaacgc  |             |            |      |
|            |            |            | acttcgaccc  |             |            |      |
|            |            |            | acaacaaaac  |             |            |      |
|            |            |            | atggtatcaa  |             |            |      |
|            |            |            | aaatcatcat  |             |            |      |
|            |            |            | acaaatccgt  |             |            |      |
|            |            |            | teggteeggg  |             |            |      |
|            |            |            | actgcaacat  |             |            |      |
| aaaccctgaa | acgtgttaaa | gaaaaactgc | aggaaaacta  | caacaacaac  | aaaaccatca | 2820 |
| aattcgctcc | gtcctccggt | ggtgacctgg | aaatcaccac  | .ccactccttc | aactgccgtg | 2880 |
|            |            |            | tgttcaacaa  |             |            |      |
|            |            |            | tcatcaacat  |             |            |      |
|            |            |            | tcacctgcaa  |             |            |      |
|            |            |            | aaaccgaaga  |             |            |      |
|            |            |            | tgtacaaata  |             |            |      |
|            |            |            | ttaccgttca  |             |            |      |
|            |            |            | gtgctatcga  |             |            |      |
|            |            |            | agacccgtgt  |             |            |      |
|            |            |            | acatgacctg  |             |            |      |
|            |            |            | tgctggaaga. |             |            |      |
|            |            |            | cctggaaaaa  |             |            |      |
|            |            |            | ctgctgtgcc  |             |            |      |
| ttgtttgccc | ctcccccgtg | ccttccttga | ccctggaagg  | tgccactccc  | actgtccttt | 3660 |
|            |            |            | gtctgagtag  |             |            |      |

```
gtggggtggg gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg 3780
atgeggtggg ctctatgggt acceaggtge tgaagaattg acceggttee teetgggeea 3840
gaaagaagca ggcacatccc ettetetgtg acacaceetg tecaegeeee tggttettag 3900
ttccaqcccc actcatagga cactcatagc tcaggagggc tccqccttca atcccacccq 3960
ctaaaqtact tqqaqcqqtc tctccctccc tcatcagccc accaaaccaa acctagcctc 4020
caagagtggg aagaaattaa agcaagatag gctattaagt gcagagggag agaaaatgcc 4080
tccaacatgt gaggaagtaa tgagagaaat catagaattt taaggccatg atttaaggcc 4140
atcatggcct taatcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct 4200
geggegageg gtateagete acteaaagge ggtaataegg ttateeaeag aateagggga 4260
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 4320
cgcgttgctg gcgtttttcc ataggctccg ccccctgac gagcatcaca aaaatcgacg 4380
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 4440
aageteeete gtgegetete etgtteegae eetgeegett aeeggataee tgteegeett 4500
tetecetteg ggaagegtgg egetttetea tageteaege tgtaggtate teagtteggt 4560
gtaggtegtt egetecaage tgggetgtgt geacgaacee eeegtteage eegacegetg 4620
cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact 4680
ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 4740
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta tctgcgctct 4800
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 4860
cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaaggatc 4920
tcaagaagat cetttgatet tttctacggg gtctgacget cagtggaacg aaaactcacg 4980
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 5040
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 5100
atgettaate agtgaggeae etateteage gatetgteta tttegtteat ceatagttge 5160
ctgactcggg ggggggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata 5220
ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5280
tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc 5340
gttqtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca 5400
aagccgccqt cccqtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 5460
tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 5520
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 5580
ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 5640
caacctatta attteccete gteaaaaata aggttateaa gtgagaaate accatgagtg 5700
acqactqaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 5760
ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 5820
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 5880
ategaatgea aceggegeag gaacactgee agegeateaa caatatttte acetgaatea 5940
ggatattett ctaatacetg gaatgetgtt tteeegggga tegeagtggt gagtaaceat 6000
qcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6060
caqtttaqtc tqaccatctc atctqtaaca tcattqqcaa cgctaccttt gccatqtttc 6120
aqaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6180
ccqacattat cqcqaqccca tttataccca tataaatcag catccatqtt gqaatttaat 6240
egeggeeteg ageaagaegt tteeegttga atatggetea taacaeceet tgtattactg 6300
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6360
catcagagat tttgagacac aacgtggctt tcccccccc cccattattg aagcatttat 6420
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6480
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 6540
                                                                  6586
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc
```

<210> 129

<211> 6505

<212> DNA

<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

tegegegitt eggigatgae ggitgaaaace tetgacacat geageteeeg gagaeggitea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattqqcca ttqcatacqt tqtatccata tcataatatq tacatttata ttqqctcatq 300 tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggqqtcatta qttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 eccgcetgge tgacegeeca acgaceceeg eccattgaeg teaataatga egtatgttee 480 cataqtaacq ccaataqqqa ctttccattq acqtcaatqq qtqqaqtatt tacqqtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tqacqqtaaa tqqcccqcct qqcattatqc ccaqtacatq accttatqqq actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatqqq cqtqqataqc qqtttqactc acggggattt ccaagtctcc accccattqa 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tettatgeat getataetgt tittggettg gggeetatae acceeegett eettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattqqtqac qatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 qqatqqqqtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet tet@cggtag cggcggaget tecacatecg agecetggte ccatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegeeacca gacataatag etgacagaet aacagaetgt teettteeat gggtetttte 1860 tgcagtcacc gtcqtcqaca cqtqtqatca qatatcqcqq ccqctctaga gatatcqcca 1920 ccatqcqtqt tcqtqqtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040 qtqttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100 acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160 cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220 ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280 aactgaccc gctgtgcgtt accctgcact gcaccaacgc tagcatcaac tgcaacgctt 2340 ccaccatcae ccaggettge ccgaaagtta acttegacee gateeegate cactactgeg 2400 ctccqqctqq ttacqctatc ctgaaatqca acaacaaaac cttctccqqt aaaqqtccqt 2460 qcaacaacqt tbccaccqtt cagtgcaccc atggtatcaa accggttqtt tccacccagc 2520 tgctgctgaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2580 acaacqttaa aaccatcatc gttcacctga acaaatccgt tgaaatcgtt tgcacccgtg 2640 ctagcgctta ctgcaacatc tccggttcca aatggaacga aaccctgaaa cgtgttaaag 2700 aaaaactgca ggaaaactac aacaacaaca aaaccatcaa attcgctccg tcctccggtg 2760 gtgacctgga aatcaccacc cactccttca actgccgtgg tgaattcttc tactgcaaca 2820 ccacccgtct gttcaacaac aacgctaccg aagacgaaac catcaccctg ccgtgccgta 2880 tcaaacagat catcaacatg tggcagggtg ttggtcgtgc tatgtacgct ccgccgatcg 2940 ctggtaacat cacctgcaaa tccaacatca ccggtctgct gctggttcgt gacggtggtg 3000

| aagacaacaa | aaccgaagaa | atchtccotc | caaataataa | taacatossa | ascasctage | 3060 |
|------------|------------|------------|------------|------------|------------|------|
|            | gtacaaatac |            |            |            |            |      |
|            | taccgttcag |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | tgctatcgaa |            |            |            |            |      |
|            | gacccgtgtt |            |            |            |            |      |
|            | catgacctgg |            |            |            |            |      |
|            | gctggaagac |            |            |            |            |      |
|            | ctggaaaaac |            |            |            |            |      |
|            | tgctgtgcct |            |            |            |            |      |
|            | cctggaaggt |            |            |            |            |      |
|            | tctgagtagg |            |            |            |            |      |
|            | ttgggaagac |            |            |            |            |      |
|            | gaagaattga |            |            |            |            |      |
|            | cacaccctgt |            |            |            |            |      |
| actcatagct | caggagggct | ccgccttcaa | tcccacccgc | taaagtactt | ggagcggtct | 3900 |
| ctccctccct | catcagccca | ccaaaccaaa | cctagcctcc | aagagtggga | agaaattaaa | 3960 |
| gcaagatagg | ctattaagtg | cagagggaga | gaaaatgcct | ccaacatgtg | aggaagtaat | 4020 |
| gagagaaatc | atagaatttt | aaggccatga | tttaaggcca | tcatggcctt | aatcttccgc | 4080 |
|            | actgactcgc |            |            |            |            |      |
|            | gtaatacggt |            |            |            |            |      |
|            | cagcaaaagg |            |            |            |            |      |
|            | cccctgacg  |            |            |            |            |      |
|            | ctataaagat |            |            |            |            |      |
|            | ctgccgctta |            |            |            |            |      |
|            | agctcacgct |            |            |            |            |      |
|            | cacgaacccc |            |            |            |            |      |
|            | aacccggtaa |            |            |            |            |      |
|            | gcgaggtatg |            |            |            |            |      |
|            | agaagaacag |            |            |            |            |      |
|            | ggtagctctt |            |            |            |            |      |
|            | cagcagatta |            |            |            |            |      |
|            | tctgacgctc |            |            |            |            |      |
|            | aggatcttca |            |            |            |            |      |
|            | tatgagtaaa |            |            |            |            |      |
|            | atctgtctat |            |            |            |            |      |
|            | cctcgtgaag |            |            |            |            |      |
|            | aaagtgaggg |            |            |            |            |      |
|            | aacttttgct |            |            |            |            |      |
|            | aactcagcaa |            |            |            |            |      |
|            | tctgccagtg |            |            |            |            |      |
|            | gaaactgcaa |            |            |            |            |      |
|            | gtaatgaagg |            |            |            |            |      |
|            | ctgcgattcc |            |            |            |            |      |
|            | ggttatcaag |            |            |            |            |      |
|            | tatgcatttc |            |            |            |            |      |
|            | tcgcatcaac |            |            |            |            |      |
|            | cgctgttaaa |            |            |            |            |      |
|            | gcgcatcaac |            |            |            |            |      |
|            | tcccggggat |            |            |            |            |      |
|            | tggtcggaag |            |            |            |            |      |
|            | cattggcaac |            |            |            |            |      |
|            | acaatcgata |            |            |            |            |      |
|            | ataaatcagc |            |            |            |            |      |
|            | tatggctcat |            |            |            |            |      |
|            | atgatatatt |            |            |            |            |      |
|            | -          | -          | -          |            |            |      |

acgtggcttt cccccccc ccattattga agcatttatc agggttattg tctcatgagc 6360 ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 6420 cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat 6480 aggcgtatca cgaggccctt tcgtc 6505

<210> 130
<211> 6466
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 130

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtct gtaageggat geegggagea gacaageeeg teagegggeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatq 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgq 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat qtcqtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac egggacegat ceagecteeg eggeegggaa eggtgeattg gaacgeggat 1020 tecceqtqcc aaqaqtqacq taaqtaccqc ctataqactc tataqqcaca cccctttqqc 1080 tottatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatqcta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecce 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040 gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100 acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160 cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catqtqqaaa aacqacatqq 2220 ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280 aactgacccc gctgtgcgtt accctgcact gcaccaacgc taqcatcaac tgcaacgctt 2340

ccaccatcac ccaggettge ccgaaagtta acttegaccc gatecegate cactactgeg 2400 ctccqqctqq ttacgctatc ctgaaatgca acaacaaaac cttctccggt aaaggtccgt 2460 gcaacaacgt ttccaccgtt cagtgcaccc atggtatcaa accggttgtt tccacccagc 2520 tgctqctqaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2580 acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcgtt tgcacccgtg 2640 ctagcgctta ctgcaacatc tccggttcca aatggaacga aaccctgaaa cgtgttaaag 2700 aaaaactgca ggaaaactac aacaacaaca aaaccatcaa attcgctccg tcctccggtg 2760 qtqacctqqa aatcaccacc cactccttca actgccgtgg tgaattcttc tactgcaaca 2820 ccgctagcct gccgtgccgt atcaaacaga tcatcaacat gtggcagggt gttggtcgtg 2880 ctatqtacqc teeqeeqate qetggtaaca teacetgcaa atecaacate aceggtetge 2940 tgctqqttcg tgacggtggt gaagacaaca aaaccgaaga aatcttccgt ccgggtggtg 3000 gtaacatgaa agacaactgg cgttccgaac tgtacaaata caaagttatc gaactgaaac 3060 cgctgggtat cgctccgacc ggtgctaagc ttaccgttca ggctcgtcag ctgctgtcct 3120 ccatcgttca gcagcagtcc aacctgctgc gtgctatcga agctcagcag cacatgctgc 3180 agctgaccgt ttggggtatc aaacagctgc agacccgtgt tctggctatc gaacgttacc 3240 tgaaagacca gcagctcgag atctggaaca acatgacctg gatggaatgg gaccgtgaaa 3300 tetecaacta cacegacace atetacegte tgetggaaga eteceagace cageaggaaa 3360 aaaacgaaaa agacctgctg gctctggact cctggaaaaa cctgtggtcc tggttcgaca 3420 tetecaactg getgtggtga ggatecagat etgetgtgee ttetagttge cagecatetg 3480 ttgtttgccc ctcccccgtg ccttccttga ccctggaagg tgccactccc actgtccttt 3540 cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct attctggggg 3600 gtggggtggg gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg 3660 atgeggtggg ctctatgggt acceaggtge tgaagaattg acceggttee teetgggeea 3720 gaaagaagca ggcacatccc cttctctgtg acacaccetg tccacgcccc tggttcttag 3780 ttccagcccc actcatagga cactcatagc tcaggagggc tccgccttca atcccacccg 3840 caagagtggg aagaaattaa agcaagatag gctattaagt gcagagggag agaaaatgcc 3960 tccaacatgt gaggaagtaa tgagagaaat catagaattt taaggccatg atttaaggcc 4020 atcategect taatetteeg etteeteget caetgacteg etgegetegg tegttegget 4080 qcqqcqagcq gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 4140 taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 4200 egegttgetg gegtttttee ataggeteeg ecceetgae gageateaca aaaategaeg 4260 ctcaaqtcaq aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 4320 aageteete gtgegetete etgtteegae eetgeegett aeeggataee tgteegeett 4380 tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt 4440 gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 4500 cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact 4560 ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 4620 cttqaaqtqq tqqcctaact acggctacac tagaagaaca gtatttggta tctgcgctct 4680 qctqaaqcca qttaccttcq qaaaaagagt tggtagctct tgatccggca aacaaaccac 4740 cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 4800 tcaagaagat cetttgatet tttetaeggg gtetgaeget cagtggaaeg aaaacteaeg 4860 ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 4920 aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 4980 atgettaate agtgaggeae etateteage gatetgteta titegtieat ceatagtige 5040 ctgactcggg gggggggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata 5100 ccaggcctga ategececat catecageca gaaagtgagg gagecaeggt tgatgagage 5160 tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc 5220 gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca 5280 aagccqccqt cccqtcaaqt cagcgtaatg ctctgccagt gttacaacca attaaccaat 5340 tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 5400 tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 5460 ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 5520 caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 5580 acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 5640

```
ggccagcat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 5700 gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt accaaacagga 5760 atcgaatgca accggcgcag gaacactgcc agcgcatcaa caatatttc acctgaatca 5820 ggatattctt ctaatacctg gaatgctgtt ttcccgggga tcgcagtggt gagtaaccat 5880 gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 5940 cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgttc 6000 agaaacaact ctggcgcatc gggcttccca tacaaatcgat agattgtcgc acctgattgc 6060 ccgacattat cgcgagcca tttataccca tataaatcag catccatgtt ggaatttaat 6120 cgcggcctcg agcaagacgt ttcccgttga atatggctca taacaccct tgtattactg 6180 tttatgtaag cagacagttt tattgtcat gatgatatat ttttatcttg tgcaatgtaa 6240 cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6360 ggggttccgc gcacatttc ccgaaaagtg ccacctgacg ttcaagaac cattattac 6420 atgacattaa cctataaaaa taggcgtat acgaggccct ttcgtc
```

<210> 131 <211> 6547 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 131 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcgqt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cqtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tettatgcat getatactgt ttttggettg gggeetatae acceeegett cettatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440

catgggetet teteoggtag eggeggaget tecacateeg agecetggte ceatgetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 ageacaatge ceaceaceae eagtgtgeeg cacaaggeeg tggeggtagg gtatgtgtet 1620 gaaaatgage gtggagattg ggetegeaeg getgaegeag atggaagaet taaggeageg 1680 geagaagaag atgeaggeag etgagttgt gtattetgat aagagteaga ggtaacteee 1740

|            |            | ggagggcagt |            | •          |            |                |
|------------|------------|------------|------------|------------|------------|----------------|
|            |            | ctgacagact |            |            |            |                |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatcgcca | 1920           |
| ccatgcgtgt | tcgtggtatc | ccgcgtaact | ggccgcagtg | gtggatgtgg | ggtatcctgg | 1980           |
| gtttctggat | gatcatcatc | tgccgtgttg | ttggtaacat | gtgggttacc | gtttactacg | 2040           |
|            |            | gctaaaacca |            |            |            |                |
|            |            | gtttgggcta |            |            |            |                |
|            |            | aacgttaccg |            |            |            |                |
|            |            | atcatctccc |            |            |            |                |
|            |            | accetgeact |            |            |            |                |
|            |            |            |            |            |            |                |
|            |            | ccgaaagtta |            |            |            |                |
|            |            | ctgaaatgca |            |            |            |                |
|            |            | cagtgcaccc |            |            |            |                |
|            |            | gctgaaaaag |            |            |            |                |
|            |            | gttcacctga |            |            |            |                |
|            |            | tccatgcgta |            |            |            |                |
|            |            | cgtcaggctt |            |            |            |                |
|            |            | gaaaaactgc |            |            |            |                |
|            |            | ggtgacctgg |            |            |            |                |
| gtgaattctt | ctactgcaac | accgctagcc | tgccgtgccg | tatcaaacag | atcatcaaca | 2940           |
| tgtggcaggg | tgttggtcgt | gctatgtacg | ctccgccgat | cgctggtaac | atcacctgca | 3000           |
| aatccaacat | caccggtctg | ctgctggttc | gtgacggtgg | tgaagacaac | aaaaccgaag | 3060           |
| aaatcttccg | tccgggtggt | ggtaacatga | aagacaactg | gcgttccgaa | ctgtacaaat | 3120           |
|            |            | ccgctgggta |            |            |            |                |
|            |            | tccatcgttc |            |            |            |                |
|            |            | cagctgaccg |            |            |            |                |
|            |            | ctgaaagacc |            |            |            |                |
|            |            | atctccaact |            |            |            |                |
|            |            | aaaaacgaaa |            |            |            |                |
|            |            | atctccaact |            |            |            |                |
|            |            | gttgtttgcc |            |            |            |                |
|            |            | tcctaataaa |            |            |            |                |
|            |            |            |            |            |            |                |
|            |            | ggtggggtgg |            |            |            |                |
|            |            | gatgcggtgg |            |            |            |                |
|            |            | agaaagaagc |            |            |            |                |
|            |            | gttccagccc |            |            |            |                |
|            |            | gctaaagtac |            |            |            |                |
|            |            | ccaagagtgg |            |            |            |                |
|            |            | ctccaacatg |            |            |            |                |
|            |            | catcatggcc |            |            |            |                |
|            |            | tgcggcgagc |            |            |            |                |
|            |            | ataacgcagg |            |            |            |                |
|            | -          | ccgcgttgct |            |            |            |                |
| cgagcatcac | aaaaatcgac | gctcaagtca | gaggtggcga | aacccgacag | gactataaag | 4380           |
| ataccaggcg | tttccccctg | gaagctccct | cgtgcgctct | cctgttccga | ccctgccgct | 4440           |
| taccggatac | ctgtccgcct | ttctcccttc | gggaagcgtg | gcgctttctc | atagctcacg | 4500           |
| ctgtaggtat | ctcagttcgg | tgtaggtcgt | tcgctccaag | ctgggctgtg | tgcacgaacc | 4560           |
|            |            | gcgccttatc |            |            |            |                |
| aagacacgac | ttatcgccac | tggcagcagc | cactggtaac | aggattagca | gagcgaggta | 4680           |
|            |            | tcttgaagtg |            |            |            |                |
|            |            | tgctgaagcc |            |            |            |                |
|            |            | ccgctggtag |            |            |            |                |
|            |            | ctcaagaaga |            |            |            |                |
|            |            | gttaagggat |            |            |            |                |
|            |            | aaaaatgaag |            |            |            |                |
|            |            |            |            |            |            | <del>-</del> - |

```
aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5100
aqaaqqtqtt qctqactcat accaggcctg aatcgcccca tcatccagcc agaaagtgag 5220
ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt tgaacttttg 5280
ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatcct tcaactcagc 5340
aaaagttega tttatteaac aaageegeeg teeegteaag teagegtaat getetgeeag 5400
tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5460
aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa 5520
ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt 5580
ccqactcqtc caacatcaat acaacctatt aatttcccct cgtcaaaaat aaggttatca 5640
aqtqaqaaat caccatqaqt gacgactgaa tccggtgaga atggcaaaag cttatgcatt 5700
tetttecaga ettgtteaac aggecageca ttacgetegt cateaaaate actegeatea 5760
accaaaccqt tattcattcq tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 5820
aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 5880
acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 5940
atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaatgctt gatggtcgga 6000
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6060
acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6120
tagattqtcq cacctgattq cccgacatta tcgcgagccc atttataccc atataaatca 6180
gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg aatatggctc 6240
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6300
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct ttccccccc 6360
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6420
atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6480
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6540
tttcgtc
```

<210> 132 <211> 6637 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 caqcttqtct qtaaqcqqat qccqqqaqca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcgggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcq qtqtqaaata ccqcacagat gcqtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 qqqqtcatta qttcataqcc catatatgga qttccqcqtt acataactta cqqtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttqqcaqtac atctacqtat taqtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020

| teccegtgee | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca               | cccctttggc | 1080 |
|------------|------------|------------|------------|--------------------------|------------|------|
|            |            |            |            | accccgctt                |            |      |
|            |            |            |            | accattattg               |            |      |
|            |            |            |            | ctctttgcca               |            |      |
|            |            |            |            | acggactctg               |            |      |
|            |            |            |            | acaacgccgt               |            |      |
|            |            |            |            | tctcgggtac               |            |      |
|            |            |            |            | agccctggtc               |            |      |
|            |            |            |            | tggaggccag               |            |      |
|            |            |            |            | tggcggtagg               |            |      |
|            |            |            |            | atggaagact               |            |      |
|            |            |            |            | aagagtcaga               |            |      |
|            |            |            |            | cagtactcgt               |            |      |
|            |            |            |            | tcctttccat               |            |      |
|            |            |            |            | ccgctctaga               |            |      |
|            |            |            |            | gtggatgtgg               |            |      |
|            |            |            |            | gtgggttacc               |            |      |
|            |            |            |            | cgcttccgac               |            |      |
|            |            |            |            | cgttccgacc               |            |      |
|            |            |            |            | catgtggaaa               |            |      |
|            |            |            |            | gtccctgaaa               |            |      |
|            |            |            |            | tagccgtaac               |            |      |
|            |            |            |            | cgctctgttc               |            |      |
|            |            |            |            | ctccgaatac               |            |      |
|            |            |            |            | taacttcgac               |            |      |
|            |            |            |            | caacaacaaa               |            |      |
|            |            |            |            | ccatggtatc               |            |      |
|            |            |            |            | agaaatcatc               |            |      |
|            |            |            |            | gaacaaatcc               |            |      |
|            |            |            |            | caaatggaac               |            |      |
|            |            |            |            | caaaaccatc               |            |      |
|            |            |            |            | caactgccgt               |            |      |
| tetactoraa | caccacccot | ctattcaaca | acaacactac | cgaagacgaa               | accatcaccc | 3000 |
|            |            |            |            | tgttggtcgt               |            |      |
|            |            |            |            | caccggtctg               |            |      |
|            |            |            |            | tccgggtggt               |            |      |
|            |            |            |            | cgaactgaaa               |            |      |
|            |            |            |            | gctgctgtcc               |            |      |
|            |            |            |            | gcacatgctg               |            |      |
|            |            |            |            | cgaacgttac               |            |      |
|            |            |            |            | ggaccgtgaa               |            |      |
|            |            |            |            | ccagcaggaa               |            |      |
|            |            |            |            | ctggttcgac               |            |      |
|            |            |            |            | ccagccatct               |            |      |
|            |            |            |            | cactgtcctt               |            |      |
|            |            |            |            | tattctgggg               |            |      |
|            |            |            |            | gcatgctggg               |            |      |
|            |            |            |            | ctcctgggcc               |            |      |
|            |            |            |            | ctggttctta               |            |      |
|            |            |            |            | aatcccaccc               |            |      |
|            |            |            |            | aacctagcct               |            |      |
|            |            |            |            |                          |            |      |
|            |            |            |            | gagaaaatgc               |            |      |
|            |            |            |            | gatttaaggc               |            |      |
|            |            |            |            | gtcgttcggc<br>gaatcagggg |            |      |
| ggcaccagcc | cacccaaagg | cygraacacy | gualicada  | 50000009999              | acaacycayy | 3240 |

```
aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 4380
ggcgtttttc cataggctcc gccccctqa cgaqcatcac aaaaatcqac gctcaagtca 4440
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4500
egtgegetet cetgtteega eeetgeeget taceggatae etgteegeet tteteeette 4560
gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt 4620
tegetecaag etgggetgtg tgeaegaace eccegtteag eccgaeeget gegeettate 4680
cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc 4740
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 4800
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaaqcc 4860
agttacette ggaaaaagag ttggtagete ttgateegge aaacaaacca cegetqqtaq 4920
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 4980
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaaqqqat 5040
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5100
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 5160
cagtgaggca cetateteag egatetgtet atttegttea tecatagttg cetgaetegg 5220
ggggggggg cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat accaggcctq 5280
aatogoocca toatocagoo agaaagtgag ggagooacgg ttgatgagag otttgttgta 5340
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcggg 5400
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5460
tecegteaag teagegtaat getetgeeag tgttacaace aattaaceaa ttetgattag 5520
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5580
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5640
atggcaagat cetggtateg gtetgegatt cegactegte caacateaat acaacetatt 5700
aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5760
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5820
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 5880
tgaqcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 5940
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 6000
totaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tgcatcatca 6060
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6120
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6180
tetggegeat egggettece atacaatega tagattgteg cacetgattg ceegacatta 6240
tegegagece atttatacce atataaatea geateeatgt tggaatttaa tegeggeete 6300
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6360
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6420
ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta tcagggttat 6480
tqtctcatqa qcqqatacat atttgaatqt atttaqaaaa ataaacaaat aggggttccq 6540
egeacattte eeegaaaagt geeacetgae gtetaagaaa eeattattat eatgacatta 6600
acctataaaa ataggcgtat cacgaggccc tttcgtc
```

```
<210> 133
<211> 6598
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 133
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
```

|             | bbbb       | tatataaata | +++-+-     | ++-        | ttaaataata | 300  |
|-------------|------------|------------|------------|------------|------------|------|
|             |            |            |            |            | ttggctcatg |      |
|             |            |            |            |            | aatcaattac |      |
|             |            |            |            |            | cggtaaatgg |      |
|             |            |            |            |            | cgtatgttcc |      |
|             |            |            |            |            | tacggtaaac |      |
|             |            |            |            |            | ttgacgtcaa |      |
|             |            |            |            |            | actttcctac |      |
| ttggcagtac  | atctacgtat | tagtcatcgc | tattaccatg | gtgatgcggt | tttggcagta | 720  |
| catcaatggg  | cgtggatagc | ggtttgactc | acggggattt | ccaagtctcc | accccattga | 780  |
| cgtcaatggg  | agtttgtttt | ggcaccaaaa | tcaacgggac | tttccaaaat | gtcgtaacaa | 840  |
|             |            |            |            |            | atataagcag |      |
|             |            |            |            |            | ttgacctcca |      |
| tagaagacac  | cgggaccgat | ccagcctccg | cggccgggaa | cggtgcattg | gaacgcggat | 1020 |
| teccegtgee  | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca | cccctttggc | 1080 |
|             |            |            |            |            | ccttatgcta |      |
| taggtgatgg. | tatagcttag | cctataggtg | tgggttattg | accattattg | accactcccc | 1200 |
| tattggtgac  | gatactttcc | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260 |
|             |            |            |            |            | tatttttaca |      |
|             |            |            |            |            | ccccgtgcc  |      |
|             |            |            |            |            | gtgttccgga |      |
|             |            |            |            |            | ccatgcctcc |      |
|             |            |            |            |            | acttaggcac |      |
|             |            |            |            |            | gtatgtgtct |      |
|             |            |            |            |            | taaggcagcg |      |
| gcagaagaag  | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
|             |            |            |            |            | tgctgccgcg |      |
| cacaccacca  | gacataatag | ctgacagact | aacagactgt | tectttecat | gggtcttttc | 1860 |
| tacaatcacc  | atcatcaaca | catataatca | gatatcgcgg | ccactctaga | gatatcgcca | 1920 |
| ccatacatat  | tegtagtate | ccacataact | gaccacagta | gtggatgtgg | ggtatcctgg | 1980 |
| atttctagat  | gatcatcatc | taccatatta | ttggtaacat | gtgggttacc | gtttactacg | 2040 |
| atattccaat  | ttggaccgac | gctaaaacca | ccctattcta | cacttccaac | accaaagcct | 2100 |
|             |            |            |            |            | gacccgaacc |      |
|             |            |            |            |            | aacgacatgg |      |
|             |            |            |            |            | ccgtgcgtta |      |
|             |            |            |            |            | tgctccttca |      |
|             |            |            |            |            | taccgtccgg |      |
| acatcottct  | actaaaaaaa | aaccotaaca | actccaacaa | ctccgaatac | atcctgatca | 2460 |
|             |            |            |            |            | ccgatcccga |      |
|             |            |            |            |            | accttctccg |      |
|             |            |            |            |            | aaaccggttg |      |
| tttccaccca  | actactacta | aacqqthccc | taactaaaaa | agaaatcatc | atccgttccg | 2700 |
| aaaacctgac  | caacaacatt | aaaaccatca | togttcacct | gaacaaatcc | gttgaaatcg | 2760 |
|             |            |            |            |            | gaaaccctga |      |
|             |            |            |            |            | aaattcgctc |      |
| catcetecaa  | taataaccta | gaaatcacca | cccactcctt | caactgccgt | ggtgaattct | 2940 |
| tctactgcaa  | caccactage | chaccatacc | gtatcaaaca | gatcatcaac | atgtggcagg | 3000 |
| atattaataa  | tactatatac | actecaceas | tractantea | catcacctac | aaatccaaca | 3060 |
| teaccooket  | actactaatt | catalcage  | atasasasas | caaaaccgaa | gaaatcttcc | 3120 |
|             |            |            |            |            | tacaaagtta |      |
|             |            |            |            |            | caggctcgtc |      |
| agetactata  | ctccatcctt | caccaccach | ccaacctoct | acatactate | gaagctcagc | 3300 |
|             |            |            |            |            | gttctggcta |      |
|             |            |            |            |            | tggatggaat |      |
|             |            |            |            |            | gactcccaga |      |
| gggaccgcga  | aacccccadc | aaaaaaata  | tractetace | chactages  | aacctgtggt | 3540 |
| cccagcayga  | aaaaaacgaa | aaayacccgc | cygecelyga | cccciggada | aaccigiggt | 2230 |

| aataattaas | catctccaac | taastataat | gaggatggag | atetestate | cettetaatt | 3600 |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            |            |      |
|            | tgttgtttgc |            |            |            |            |      |
|            | ttcctaataa |            |            |            |            |      |
|            | gggtggggtg |            |            |            |            |      |
| ggcatgctgg | ggatgcggtg | ggetetatgg | gracecaggr | gccgaagaac | tgaceeggtt | 3840 |
|            | cagaaagaag |            |            |            |            |      |
| cctggttctt | agttccagcc | ccactcatag | gacactcata | gctcaggagg | gctccgcctt | 3960 |
|            | cgctaaagta |            |            |            |            |      |
|            | tccaagagtg |            |            |            |            |      |
| agagaaaatg | cctccaacat | gtgaggaagt | aatgagagaa | atcatagaat | tttaaggcca | 4140 |
| tgatttaagg | ccatcatggc | cttaatcttc | cgcttcctcg | ctcactgact | cgctgcgctc | 4200 |
| ggtcgttcgg | ctgcggcgag | cggtatcagc | tcactcaaag | gcggtaatac | ggttatccac | 4260 |
| agaatcaggg | gataacgcag | gaaagaacat | gtgagcaaaa | ggccagcaaa | aggccaggaa | 4320 |
|            | gccgcgttgc |            |            |            |            |      |
|            | cgctcaagtc |            |            |            |            |      |
|            | ggaagctccc |            |            |            |            |      |
|            | tttctccctt |            |            |            |            |      |
|            | gtgtaggtcg |            |            |            |            |      |
|            | tgcgccttat |            |            |            |            |      |
|            | ctggcagcag |            |            |            |            |      |
|            | ttcttgaagt |            |            |            |            |      |
|            | ctgctgaagc |            |            |            |            |      |
|            | accgctggta |            |            |            |            |      |
|            | tctcaagaag |            |            |            |            |      |
|            | cgttaaggga |            |            |            |            |      |
| cyadaaccca | taaaaatgaa | attttaaata | astotasant | atatatgagt | asacttooto | 5100 |
|            | caatgcttaa |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | gcctgactcg |            |            |            |            |      |
|            | taccaggcct |            |            |            |            |      |
| gttgatgaga | gctttgttgt | aggtggacca | greggreate | ttgaacttt  | getttgecae | 5340 |
| ggaacggtct | gcgttgtcgg | gaagatgcgt | gatetgatee | ttcaactcag | caaaagttcg | 5400 |
|            | caaagccgcc |            |            |            |            |      |
|            | attctgatta |            |            |            |            |      |
| atatcaggat | tatcaatacc | atatttttga | aaaagccgtt | tctgtaatga | aggagaaaac | 5580 |
| tcaccgaggc | agttccatag | gatggcaaga | tectggtate | ggtctgcgat | tccgactcgt | 5640 |
|            | tacaacctat |            |            |            |            |      |
|            | tgacgactga |            |            |            |            |      |
| acttgttcaa | caggccagcc | attacgctcg | tcatcaaaat | cactcgcatc | aaccaaaccg | 5820 |
| ttattcattc | gtgattgcgc | ctgagcgaga | cgaaatacgc | gatcgctgtt | aaaaggacaa | 5880 |
| ttacaaacag | gaatcgaatg | caaccggcgc | aggaacactg | ccagcgcatc | aacaatattt | 5940 |
|            | caggatattc |            |            |            |            |      |
| gtgagtaaco | atgcatcatc | aggagtacgg | ataaaatgct | tgatggtcgg | aagaggcata | 6060 |
| aattccgtca | gccagtttag | tctgaccatc | tcatctgtaa | catcattggc | aacgctacct | 6120 |
| ttgccatgtt | tcagaaacaa | ctctggcgca | tegggettee | catacaatcg | atagattgtc | 6180 |
| gcacctgatt | gcccgacatt | atcgcgagcc | catttatacc | catataaatc | agcatccatg | 6240 |
| ttggaattta | atcgcggcct | cgagcaagac | gtttcccgtt | gaatatggct | cataacaccc | 6300 |
| cttgtattac | tgtttatgta | agcagacagt | tttattgttc | atgatgatat | atttttatct | 6360 |
| tatacaatat | aacatcagag | attttgagac | acaacgtggc | tttcccccc  | ccccattat  | 6420 |
| tgaagcattt | atcagggtta | ttqtctcatq | ageggataca | tatttqaatq | tatttagaaa | 6480 |
|            | taggggttcc |            |            |            |            |      |
|            | tcatgacatt |            |            |            |            | 6598 |
| accactacta | gacacc     |            |            |            |            |      |

<211> 6679 <212> DNA

<213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 134 tegegegett eggegatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta qttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cocgeotgge tgacegeeca acgaceceeg eccattgacg teaataatga egtatgttee 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtacege etatagaete tataggeaea eccetttgge 1080 tettatgcat getatactgt ttttggettg gggeetatac accecegett cettatgeta 1140 taggtgatgg tatagcttag cetataggtg tgggttattg accattattg accactecce 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeg agecetggte ecatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcqqtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt teetttecat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040 gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100 acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160 cgcaggaaat cqttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgq 2220 ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280 aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagccgtaac tgctccttca 2340 acaccaccac egaaateegt gacaaaaaac agcagggtta egetetgtte tacegteegg 2400 acategitet getgaaagaa aacegtaaca actecaacaa eteegaatac ateetgatea 2460 actgcaacgc ttccaccatc acccaggett gcccgaaagt taacttcgac ccgatcccga 2520 tocactactq cqctccqqct ggttacgcta tcctqaaatq caacaacaaa accttctccq 2580 gtaaaggtcc gtgcaacaac gtttccaccg ttcagtgcac ccatggtatc aaaccggttg 2640 tttccaccca gctgctgctg aacggttccc tggctgaaaa agaaatcatc atccgttccg 2700 aaaacctgac cgacaacgtt aaaaccatca tcgttcacct gaacaaatcc gttgaaatcg 2760 tttgcacceg tccgaacaac aacacccgta aatccatgcg tatcggtccg ggtcagacct 2820

|            | cggtgacatc |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | cgaaaccctg |            |            |            |            |      |
|            | caaatteget |            |            |            |            |      |
|            | tggtgaattc |            |            |            |            |      |
|            | catgtggcag |            |            |            |            |      |
|            | caaatccaac |            |            |            |            |      |
|            | agaaatcttc |            |            |            |            |      |
|            | atacaaagtt |            |            |            |            |      |
| agcttaccgt | tcaggctcgt | cagctgctgt | cctccatcgt | tcagcagcag | tccaacctgc | 3360 |
|            | cgaagctcag |            |            |            |            |      |
|            | tgttctggct |            |            |            |            |      |
|            | ctggatggaa |            |            |            |            |      |
|            | agactcccag |            |            |            |            |      |
|            | aaacctgtgg |            |            |            |            |      |
|            | gccttctagt |            |            |            |            |      |
| tgaccctgga | aggtgccact | cccactgtcc | tttcctaata | aaatgaggaa | attgcatcgc | 3780 |
| attgtctgag | taggtgtcat | tctattctgg | ggggtggggt | ggggcaggac | agcaaggggg | 3840 |
|            | agacaatagc |            |            |            |            |      |
| tgctgaagaa | ttgacccggt | tectectggg | ccagaaagaa | gcaggcacat | ccccttctct | 3960 |
| gtgacacacc | ctgtccacgc | ccctggttct | tagttccagc | cccactcata | ggacactcat | 4020 |
| agctcaggag | ggctccgcct | tcaatcccac | ccgctaaagt | acttggagcg | gtctctccct | 4080 |
| ccctcatcag | cccaccaaac | caaacctagc | ctccaagagt | gggaagaaat | taaagcaaga | 4140 |
| taggctatta | agtgcagagg | gagagaaaat | gcctccaaca | tgtgaggaag | taatgagaga | 4200 |
| aatcatagaa | ttttaaggcc | atgatttaag | gccatcatgg | ccttaatctt | ccgcttcctc | 4260 |
| gctcactgac | tegetgeget | cggtcgttcg | gctgcggcga | gcggtatcag | ctcactcaaa | 4320 |
| ggcggtaata | cggttatcca | cagaatcagg | ggataacgca | ggaaagaaca | tgtgagcaaa | 4380 |
| agġccagcaa | aaggccagga | accgtaaaaa | ggccgcgttg | ctggcgtttt | tccataggct | 4440 |
| ccgccccct  | gacgagcatc | acaaaaatcg | acgctcaagt | cagaggtggc | gaaacccgac | 4500 |
|            | agataccagg |            |            |            |            |      |
|            | cttaccggat |            |            |            |            |      |
| tcatagctca | cgctgtaggt | atctcagttc | ggtgtaggtc | gttcgctcca | agctgggctg | 4680 |
| tgtgcacgaa | cccccgttc  | agcccgaccg | ctgcgcctta | tccggtaact | atcgtcttga | 4740 |
| gtccaacccg | gtaagacacg | acttatcgcc | actggcagca | gccactggta | acaggattag | 4800 |
| cagagcgagg | tatgtaggcg | gtgctacaga | gttcttgaag | tggtggccta | actacggcta | 4860 |
| cactagaaga | acagtatttg | gtatctgcgc | tctgctgaag | ccagttacct | tcggaaaaag | 4920 |
|            | tcttgatccg |            |            |            |            |      |
| caagcagcag | attacgcgca | gaaaaaaagg | atctcaagaa | gatcctttga | tcttttctac | 5040 |
|            | gctcagtgga |            |            |            |            |      |
| aaaaaggatc | ttcacctaga | tccttttaaa | ttaaaaatga | agttttaaat | caatctaaag | 5160 |
|            | taaacttggt |            |            |            |            |      |
|            | ctatttcgtt |            |            |            |            |      |
| tctgcctcgt | gaagaaggtg | ttgctgactc | ataccaggcc | tgaatcgccc | catcatccag | 5340 |
| ccagaaagtg | agggagccac | ggttgatgag | agctttgttg | taggtggacc | agttggtgat | 5400 |
| tttgaacttt | tgctttgcca | cggaacggtc | tgcgttgtcg | ggaagatgcg | tgatctgatc | 5460 |
| cttcaactca | gcaaaagttc | gatttattca | acaaagccgc | cgtcccgtca | agtcagcgta | 5520 |
| atgctctgcc | agtgttacaa | ccaattaacc | aattctgatt | agaaaaactc | atcgagcatc | 5580 |
|            | gcaatttatt |            |            |            |            |      |
|            | aaggagaaaa |            |            |            |            |      |
| cggtctgcga | ttccgactcg | tccaacatca | atacaaccta | ttaatttccc | ctcgtcaaaa | 5760 |
|            | caagtgagaa |            |            |            |            |      |
|            | tttctttcca |            |            |            |            |      |
|            | caaccaaacc |            |            |            |            |      |
|            | taaaaggaca |            |            |            |            |      |
|            | caacaatatt |            |            |            |            |      |
| gttttcccgg | ggatcgcagt | ggtgagtaac | catgcatcat | caggagtacg | gataaaatgc | 6120 |

```
ttgatggteg gaagaggcat aaatteegte agecagttta gtetgaceat eteatetgta 6180 acateattgg caaegetace tttgccatgt tteagaaaca actetggege ategggette 6240 ceatacaate gatagattgt egeacetgat tgccegacat tategegage ecatttatac 6300 ceatataaat cagcatecat gttggaattt aategeggee tegageaaga egttteeegt 6360 tgaatatgge teataacace eettgtatta etgtttatgt aageagacag ttttattgtt 6420 catgatgata tattttate ttgtgeaatg taacateaga gattttgaga eacaaegtgg 6480 ettteeece eeeecatta ttgaageatt tateagggtt attgtetat gageggatac 6540 atatttgaat gtatttaga aaataaacaa ataggggtte egegeacatt teeeegaaaa 6600 gtgeeacetg aegtetaaga aaceattatt ateatgacat taacetataa aaataggegt 6660 ateaegagge eeettteegte
```

<210> 135
<211> 6646
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 135

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tqcccacttq qcaqtacatc aaqtqtatca tatqccaaqt acqccccta ttqacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tottatgcat gotatactgt ttttggcttg gggcctatac acccccgctt cottatgcta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac totgtcottc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgo ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aaqagtcaqa ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegecacca gacataatag etgacagact aacagactgt teetttecat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980

| gtttctggat  | gatcatcatc  | taccatatta  | ttggtaacat | gtgggttacc  | gtttactacg  | 2040 |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| gtgttccggt  | ttggaccgac  | gctaaaacca  | ccctattcta | cqcttccqac  | accaaagcct  | 2100 |
| acgaccotga  | agttcacaac  | atttagacta  | cccacgcttg | cqttccqacc  | qacccgaacc  | 2160 |
| cacaagaaat  | cattctagaa  | aacqttaccq  | aaaacttcaa | catqtqqaaa  | aacqacatqq  | 2220 |
| ttgaccagat  | gcacgaagac  | atcatctccc  | tgtgggacca | gtccctgaaa  | ccqtqcqtta  | 2280 |
| aactgacccc  | actatacatt  | accetacact  | gcaccaacgc | taccttcaaa  | aacaacqtta  | 2340 |
| ccaacgacat  | gaacaaagaa  | arccgtaact  | gctccttcgc | atocatcaac  | tocaacoctt  | 2400 |
| ccaccatcac  | ccagacttac  | ccgaaagtta  | acttcgaccc | gatecegate  | cactactaca  | 2460 |
| ctccaactaa  | tracgetate  | chgaaatgca  | acaacaaaac | cttctccqqt  | aaaggtccgt  | 2520 |
| gcaacaacgt  | thecaccatt  | cagtgcaccc  | atggtatcaa | accogttott  | tccacccagc  | 2580 |
| tactactasa  | caattccta   | actaaaaaaa  | aaatcatcat | ccattccaaa  | aacctgaccg  | 2640 |
| acaacottaa  | aaccatcatc  | gttcacctga  | acaaatccgt | tgaaatcgtt  | tacacccatc  | 2700 |
| ccaacaacaa  | cacccgtaaa  | tccatgcgta  | teggteeggg | tcagaccttc  | tacactacca  | 2760 |
| ataacatcat  | caataacata  | catcagactt  | actgcaacat | ctccaattcc  | aaatqqaacq  | 2820 |
| aaaccctgaa  | acgtgttaaa  | gaaaaactgc  | aggaaaacta | caacaacaac  | aaaaccatca  | 2880 |
| aattccctgaa | atcetecast  | gataacctga  | aaatcaccac | ccactccttc  | aactgccgtg  | 2940 |
| gtgaattett  | chactecase  | accacccatc  | tgttcaacaa | caacgctacc  | gaagacgaaa  | 3000 |
| ccatcaccct  | accataccat  | atcaaacaga  | tcatcaacat | ataacaaaat  | attaatcata  | 3060 |
| ctatcacccc  | tergergate  | actaataaca  | tcacctgcaa | atccaacatc  | accoatctac  | 3120 |
| tactacttca  | taecaataat  | geeggeaaca  | aaaccgaaga | aatcttccgt  | ccaaataata  | 3180 |
| atagataga   | agacagacaga | cattecase   | tgtacaaata | caaagttatc  | gaactgaaac  | 3240 |
| graacargaa  | cactacacag  | agtactagaac | ttaccgttca | aactcatcaa  | ctactatect  | 3300 |
| cgctgggtat  | geocegace   | aacctactac  | gtgctatcga | agetegeeag  | cacatactac  | 3360 |
| ccatcgttca  | ttagggtate  | aaccegeege  | agacccgtgt | tctqqctatc  | gaacgttacc  | 3420 |
| agetgacege  | ccggggcacc  | atctggagg   | acatgacctg | datadaatad  | gaccgtgaaa  | 3480 |
| tgaaagacca  | geagetegag  | atctggaaca  | tgctggaaga | ctcccacacc  | caccaccaca  | 3540 |
| ccccaacta   | caccyacacc  | acceaccyce  | cctggaaaaa | cctataatcc  | tagttaggaaa | 3600 |
| aaaacgaaaa  | agaeeeegeeg | gastagagat  | ctgctgtgcc | ttctagttgc  | cagccatcto  | 3660 |
| terecaacig  | geegeggega  | ggattcagat  | ccctggaagg | toccastocc  | actotecttt  | 3720 |
| ctgtttgeec  | tanaganatt  | ggatgggatt  | gtctgagtag | atatcattct  | attctogggg  | 3780 |
| cctaataaa   | gaggaaacc   | aaaaaaaaa   | attgggaaga | caataggagg  | catactaga   | 3840 |
| gragagraga  | graygarage  | aagggggagg  | tgaagaattg | accognition | tectagacca  | 3900 |
| acgeggeggg  | gggagateg   | cttctctata  | acacaccctg | tccacgcccc  | taattettaa  | 3960 |
| gaaagaagca  | agtactacce  | cactcatage  | tcaggagggc | tecaegette  | atcccacccg  | 4020 |
| ghaaghagh   | tagaaaaata  | tetecetece  | tcatcagccc | accaaaccaa  | acctagecte  | 4080 |
| caaagtact   | aagaaattaa  | agraagatag  | gctattaagt | acsassassa  | agaaaatgcc  | 4140 |
| tecasestet  | gaggaactaa  | tagaagaaaat | catagaattt | taaggccatg  | atttaaggc   | 4200 |
| atcategest  | taatetteen  | cttcctcact  | cactgactcg | ctacactcaa  | tcattcaget  | 4260 |
| accacagaca  | chatcacete  | actcaaaggc  | ggtaatacgg | ttatccacag  | aat.cagggga | 4320 |
| taacacaaa   | agagagatat  | acceanage   | ccagcaaaag | accaddaacc  | ghaaaaaggc  | 4380 |
| caacgcagga  | acattttcc   | ataggaaaagg | ccccctgac  | gagcatcaca  | aaaatcqacq  | 4440 |
| ctcaactcac  | acataccaa   | acaggeeeeg  | actataaaga | taccaggest  | thereestag  | 4500 |
| ageteete    | atacaststs  | ctattccasc  | cctgccgctt | accogatace  | tatecacett  | 4560 |
| tataaattaa  | gegegeeeee  | castttetea  | tagctcacgc | tatagatata  | tcagttcggt  | 4620 |
| ghagghagh   | ggaagcgcgg  | tagactatat  | gcacgaaccc | cccattcac   | ccaaccacta  | 4680 |
| graggreger  | gataactate  | atattaaata  | caacccggta | agacacgach  | rategeeact  | 4740 |
| geeccace    | ggtaattate  | gcctcgagcc  | agcgaggtat | atacacgata  | ctacagagtt  | 4800 |
| ggcagcagcc  | baggetaaca  | accageag    | tagaagaaca | grafftggta  | tetocactet  | 4860 |
| actassaces  | attacettee  | deadacac    | tggtagctct | toatecooca  | aacaaaccac  | 4920 |
| gergaageca  | gulaculucg  | ttatttacee  | gcagcagatt | acacacagea  | aaaaaaaatc  | 4980 |
| teaacaaca   | ggtggtttt   | tttctaccc   | gtctgacgct | cartrosaco  | aaaactcaco  | 5040 |
| ttaagaagat  | ttaataataa  | gattatcasa  | aaggatcttc | acctacatec  | ttrtaaatta  | 5100 |
| aaaataaaat  | tttasstass  | totasantat  | atatgagtaa | acttootcto  | acagthacca  | 5160 |
| addatyddgt  | aghgagggaa  | ctatctcacc  | gatctgtcta | tttcattcat  | ccataottoc  | 5220 |
| aryerraare  | agryayycac  | actasaatet  | gcctcgtgaa | gaaggtgttg  | chaactcata  | 5280 |
| cryacteygg  | 222222222   | 200202200   | Jeeeegegaa | 2~~22~2     |             |      |

```
ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5340
tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc 5400
gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca 5460
aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 5520
tctqattaqa aaaactcatc qagcatcaaa tqaaactgca atttattcat atcaggatta 5580
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 5640
ttccataqqa tqqcaaqatc ctqqtatcqq tctqcqattc cqactcqtcc aacatcaata 5700
caacctatta atttcccctc qtcaaaaata aggttatcaa qtqaqaaatc accatqaqtq 5760
acqactqaat ccqqtgagaa tgqcaaaagc ttatgcattt ctttccagac ttgttcaaca 5820
ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 5880
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 5940
atcqaatqca accqqcqcaq qaacactqcc aqcqcatcaa caatattttc acctqaatca 6000
ggatattett etaatacetg gaatgetgtt tteeegggga tegeagtggt gagtaaceat 6060
qcatcatcaq qaqtacqqat aaaatqcttq atqqtcqqaa qaqqcataaa ttccqtcaqc 6120
caqtttaqtc tqaccatctc atctqtaaca tcattqqcaa cgctaccttt gccatqtttc 6180
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattqc 6240
ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt ggaatttaat 6300
egeggeeteg ageaagaegt tteeegttga atatggetea taacaceeet tgtattactg 6360
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6420
catcagagat tittgagacac aacgiggett teceeecee eccattatig aageatitat 6480
caqqqttatt qtctcatqaq cqqatacata tttqaatqta tttagaaaaa taaacaaata 6540
qqqqttcqqc qcacatttcc ccqaaaaqtq ccacctqacq tctaaqaaac cattattatc 6600
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc
```

```
<210> 136
<211> 6565
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 136
```

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcitatet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120 ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 qqqqtcatta qttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cocqcetqqc tqaccqccca acqacccccq cccattqacq tcaataatqa cqtatqttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tqcccacttq qcagtacatc aaqtqtatca tatqccaagt acqccccta ttqacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttqqcaqtac atctacqtat tagtcatcgc tattaccatg qtqatgcggt tttggcaqta 720 catcaatggg cgtggatage ggtttgacte acggggattt ccaagtetee accecattga 780 cqtcaatqqq aqtttqtttt qqcaccaaaa tcaacqggac tttccaaaat qtcqtaacaa 840 ctccqcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 ageteqttta qtgaaccqtc agategeetg gagaegeeat ccacqctgtt ttgacctcca 960 tagaagacac cqqqaccqat ccaqcctccq cqqccqqqaa cqqtqcattq qaacqcqqat 1020 tecceqtgce aagagtgacg taagtacege etatagacte tataggcaca eccetttggc 1080 tettatgeat getatactqt ttttggettg gggeetatac acceecqett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecce 1200

| Eatleghgac galactitice altactaate catacatigg cictitiges agactatette 1260 tatlegotat altgecaata tottgictotte agagactiga acggactotty tattitata 1220 ggatggggtc ceatitatia titacaaati cacatataa acaacgccgt cecccgigcc 1380 egaggggggggggggggggggggggggggggggggggg   |            |            |            |            |            |            |      |
|--|------------|------------|------------|------------|------------|------------|------|
| ggatgggtc tchtcqgdag cggqggggagt tcacqaagcag tcccqggatc tggggggcgggggggggggggggggggggggggg   |            |            |            |            |            |            |      |
| cgcagtttt attaaacata gctgtggate tecacgcgaa tetegggtac gtgttectg catgggetet tectocggtag eggeggagt tecacaccac accaccacc cacquetet caccacaccac cacquetetg caccacaccac gagetgtggg cagatettg getectac ggaaagaaga tagaggaagac tagaggagaagac tagaggaagac tagaggaagac tagaggaagaagac ggaaagaagac ggagaagaagac tagaggaagac gagaagaagac tagaggaagac gagagaagac gagagaga  | tattggctat | atgccaatac | tctgtccttc | agagactgac | acggactctg | tatttttaca | 1320 |
| actgggctcat ggtcgetcgg cagctccttg ctcctaacag ggcctggtc catggccaca 1500 agcagcatatg ggtcgctcag cagctccttg ctcctaacag tggaggcag acttaggcac 1560 agcacaatgc coaccacaca cagttgtgccg cacaaggccg tggaggtagg gtatgtgtct 1620 gaaaatgag gtgaggagd ggtcgcacag ggtaggagd ggtagggcag ggaggagg ggaggaggaggaggaggaggaggaggag  |            |            |            |            |            |            |      |
| actgggctcat ggtcgetcgg cagctccttg ctcctaacag ggcctggtc catggccaca 1500 agcagcatatg ggtcgctcag cagctccttg ctcctaacag tggaggcag acttaggcac 1560 agcacaatgc coaccacaca cagttgtgccg cacaaggccg tggaggtagg gtatgtgtct 1620 gaaaatgag gtgaggagd ggtcgcacag ggtaggagd ggtagggcag ggaggagg ggaggaggaggaggaggaggaggaggag  | cgcagttttt | attaaacata | gcgtgggatc | tccacgcgaa | tctcgggtac | gtgttccgga | 1440 |
| agoggotca gytogotog cagotoctty coctaacy tygaggocay acttaggoa 1560 gaaaatgag cocacacca cagtytycc cacaagoccy tygggytagy gtatytyct 1620 gaaaatgag gygagatty gytogocag gytagacya atgaaagat taaggoagy 1680 gcagaaagaa atgaaggaa ctgagtyty gtattotya gagtactogt tygtgegytagy gytaactcc 1740 gtatgogytog tytaaacga cygagagaa gagaagaat taaggaaga 1860 gcagcacca gacataatag ctgacagac acaagacty tocttocat gygtotttc 1860 tygagtacac gytogogaa cytygatac gagaagaat cacagacty toctygagagagat gytaatcyag 1890 gtttctggat totyggatac cogotyaact gygagagaga gygaatcyg gytaatcyg 1890 gtttctggat gatcacaac gyttygatac cocgtotty tyggatygag gytaatcyg 1890 gtttctggat gytacacaa gyttyggata cocgatyg tyggatygag gytaatcyg 1890 gytttctggat gytacacaac gyttyggata cocgacyty gygatygag gytaatcyg 1890 gytttctggat gytacacaac gyttyggata cocgacyty gygatygag gytaatcyg 1890 gytttctggat gyttacaaac gyttyggyta cocgacyty gygatygag gytaatcyg 1890 gyttygaga agttacacaa gyttyggata cocgacyty gygatygag gytaaccyg 1890 gyttyggyg agyttocyg agytaccacaac gyttygagaca acgacyaacyg 2220 tygacagagaa cyttyggat acctygacac tyggaacacyg gyttacgac gyttocygac gaccygaac gyctygagagaa acgacyaacyg 2220 tygacagaa gygtaccyga cyttygagaa aacgacaacyg 2220 tygacagaa gyttocygaa acctygacac gygacaca gyttygagaa aacgacatyg 2220 tygacagaa gyttyggat acctygacac gygacacacacyt gygaacacacacyt gygaacacacacacacacacacacacacacacacacacac   |            |            |            |            |            |            |      |
| agacacaacga         caccaccacca         cagtgtgccg         gctgaggattg         ggctgccacg         gtgaggattg         1680           gaaaatgaaga         gtgaggattg         ggctgccacg         gtgagtccg         1740           gttgggtgc         gttaacggt         ggaggagatg         tjagttttg         gtattctgat         aagaatcacg         ggtaacccc         1740           gttgcggtcac         gttaacggt         ggaggagatg         gattactcggt         gggttactc         1800           gttttctggat         gattactact         cactggttt         tgtgggttac         ggtttactgg         ggttactcgg         ggttactcgg         1920           gttttctggat         gattactact         gctcgtgacg         gctaacacg         ggttactcgg         ggttactcgg         ggttactcgg         1920           gtttctggat         gattactact         gcttgggattac         ggttactgg         ggttactgg         ggttactcgg         1920           gtttccggat         gattaccacac         gtttggattac         gcttccttacg         ggattactg         2946           gttaccacaga         gattactacac         gtttactactg         gattactacac         accacactgtacacacacacacacacacacacacacacac   |            |            |            |            |            |            |      |
| gaaaatgagg gtgagaattg ggctegcag gtgatctg atggaagact taaggcagg 1680 gcagaagaag atggaagacg ctgagttgt gtatctgat aaggtcagg gtgaactcc 1740 gttgcggtgc tgttaacgg gaaggcagt gtatctgag cagtactcgt tgttacacgg gaataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 gcagtcacc gtgtcgcac cgtgtgatca cggctgacgt ggcagtagg gtgaatgtgg gtgatcctgg 1920 ccatgcgtgt tcgtggtatc cgcgtgatca ggccgactg gtgagatggg gtgatctgg 1920 gtgttctgga gataatcact tgccgtgttg ttggacacat ggccgcactg gtggatgagg gtgatctgg 1920 gtgttctgga accaaagact 2100 acgaccgtga agttcacac gctttgggcac gctaaacaca cctgttctg gcttccgac accaaagact 2100 acgaccgtga agttcacaac gtttgggcta cccacactcac cgttctgga aacgacagg gtaacacaa accaaagacat gatcacaaa accataccc tgtggggaca gaccaaaaga ccgttctggaa aacgataccc gcacacacgt cacacaacgta cacaaagacat gacaaaagaa atccgtaact gcccctacac gaccaaaaga ccgtgagtta acctgaact gcccacaacga gaacaaaga atccgtaact gcccctacac gaccaaaaga accacaaagacacg ctacacacacgta gacaaaaga atccgtaact gccccttcgc atgcacacacgt cactactcg 2400 ccacacacac ccaggcttg ccgaaaagaa aacaacaca accacacacacacacac  |            |            |            |            |            |            |      |
| geagaagaaa atcaagagaa ctaagttgtt gtaatcctga aagaactaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactagt ggtagtcaga ggcagtcagca geataataag ctgacagaactagaa aacagaactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtgaac cgtgtgatcac gtagtagtag ggtaatcctg 1920 catagcgtgt tcgtggaca cgtgtgatcac ggcagtagtg gggatgtgg ggtatcctgg 1980 tttctggat gatactact tgcggtgttg ttggtaacat gtgggtgtac ggtaactagt gttggacac gttagaacaa gtttacaaac gtttgggca cccaagcgt gggatgtgg ggtatcctgg 1980 acgacggtga agttacaaac gttttgggca cccaacgct ggttccgac gatcaagacc 2040 gggaggaaat cgttctggaa aacgttaccg aaaactaca catgtgggaa aacgacagaa cccaagcgt gacagaagaa atcatctcc tgtgggacca gtccatgaa aacgacatgg 2220 ttgacacaag gacagaagaa atcatctccc tgtgggacca gtccatgaaa acgacatgg 2220 ttgacacaag gacaaagaa atcatcacc gcacaacgc taccttcaaa aacgacatga 2240 cccacatcac ccaggcttgc ccgaaagta gctccttcgc atgacacac gctcctggta 2240 ccacacacca ccargcttgc ccgaaagta atcaccac cacacacca tcctccaa aacacacac  |            |            |            |            |            |            |      |
| getgecacea gacataadag etgacagact acagacett tectetteca gggetettete 1860 tegagtace gtegtegace ctgstgatea gatategge cegetetaga gatategge 1820 ceatgegtgt tegtgegaca cegtgtagate gatategge gegetettegaga gatategge 1820 gtttettegat gatateggea getgacagatg gtggatgtgg ggtatectgg 1920 gtttettegat gatateggea getaaaaca cectgttetg gegetegaag gttacacag 2040 gtgttetggat agtateacaa gtttgggata cectgttetg gegtteegac gacaaaagac 2040 gtgagagaat gattacaaac gtttgggata cectgttetg gegtteegac gacaaagac 2100 acgaceggaa agttacaaac gtttgggata accatgacegg gegtggaaa acgacegaac 2160 acgacegaat gacaaagaa atcatetece tgtgggaca gtcectgaaa acgacegaca 2220 ttgaccagat gacaaagaa atcatetece tgtgggaca gtcectgaaa acgacegaca 2220 ttgaccagat gacaaagaa atcatetece tgtgggaca gtcectgaaa acgacegta 2220 ttgacacaa gegttgegt accetgacac gtcecttega gtgatacaa acaacaacgt 2220 acacaacaca gaacaagaa atcateacaca ctgaaagac atcatetece gatecettega gtgatacaa acaacaacgt 2220 acacaacaca gaacaacga accagaagaa accatgaceg 2460 acaacaacgt ttcacegtt cagtacca acaacaacaca accagetta 2400 acaacagttaa agcateact ggtsacacca atggtateaa accagttaca accaacacgt accaacacgtta 2400 acaacagttaa agcateact gtgaaacaca gtgtsaccac gatecegac 2460 acaacagttaa agcateact gtgaaacaca gaaacacgtta 2400 acaacagttaa agcateact gtgaaacaca accaacacaca accacacaca acca |            |            |            |            |            |            |      |
| gegecacea gacataataa etgacagac aacaagac getstetca gatategee 1920 ccategogty tegtsgatac cegetstact gacaceget gtggatetag gatategee 1920 gtttetegat gatatacate becegtstit tegtsaacat gtggstace gtttactaag 2040 gtgtteeggt ttggaceae gttaggaca cectagtete ggtteegae gtttactaag 2040 gtggtteeggt ttggaceae gttaggaca cectagtete ggtteegae gtttactaag 2040 gtgacagaaat egttetggaa aacatactee tegtggtta cecaegetty egtteegae accaaagee 2160 cgcaagaaat cgttetggaa aacatetee tegtggaca gacaceaegg 2220 ttgaccagat gacagaaga atcatetee tygggaca gteectgaaa aacgacatgg 2220 ttgaccagat gacagaagac atcatetee tygggaca gteectgaaa aacgacatgg 2220 tegaccagaat gacagaagaa atcetteee tygggaca gteectgaaa aacgacatgg 2220 ccaacacace cacggettye cegaaagta gedeceeae tacetteaa aacaacagta 2340 ccaacacaca cacagettee cegaaagta accetagace gatecegate dacetacege 2860 ccaccateae cacagettee cegaaagta accetagace gatecegate accetacege 2860 ccaccateae cacagettee cegaaagta accetagaac cegatecegate 2820 gcaacaacgt ttecacegt cagaaagaa aacaacaaca accaacaaca accagaca aacaacaca aacaacaca accagacat aacaacacaca aacaacaca accagettee tygaaatege 2820 ctaggotta ctgcaacate tecggtteea aagagacaac cegtteegga aacacacaca aacaacaca aacacacaca aacacacacaca aacacacaca aacacacacaca aacacacaca aacacacacaca aacacacaca aacacacaca aacacacaca aacacacaca aacacacaca aac |            |            |            |            |            |            |      |
| EgoagtoaccEtcgtgatacccgtstgtatacccgtstgtatacccgtstgtataccgatatcgatggtstgtstgtaggtstatctactgytogtataccgytogtatttaccgytogtattaccgytogtattaccgytogtattaccgytogtattaccgytotaccag <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>   |            |            |            |            |            |            |      |
| ccatgoggts togtggtate cogogtaact ggcggaagt gtggatgtgg ggtatectgg 1980 gtttettggat gateateate tpccgtgttg ttggtaacat gtgggttace gtttactaceg 2040 gtgttccggt ttggacogac gctaaaacca coctgttctg cgcttccgac accaaagcdt 2100 acgacogtga agttcacaac gtttgggcta cacaacgctg cgttccgac gaccagaacc 2160 cgcaggaaat cgttctggaa aacgttaccg aaacattcac catgtgggaca accgaacgtg 2280 acctgaacgt gacaagaag atccttcact tgtgggacaa gtccctgaaa ccgttgggtt 2280 acctgacctg gacaagaaga atccgtaact ggtccttcgc gtcactacaa ccgacagtgt 2280 ccacaacacac gacaagaaga atccgtaact gcccacacgc taccttcaa aacaacgtta 2340 ccacacacact gacaagaga atccgtaact gcccacacgc taccttcaa acaaggtcgt 2280 gcaacaacgt ttccaccgtt cogyaaagtta acttcgaccc gatcccgatc cactactgg 2460 ctccgggtgg ttacgetgt cogyaaagata acttcgaccc atggtacca accggtbgt tccaccaga 2880 gcgtggdata aaccatcata gttcacctga aaacaacaa aaccatcaa agttcacacac cactacttca cagtgacacc aaggaacaca atccgtscgaa aacctgaaa accgtgtaaa aaccagaa aacaacacaa aacaacacaa aacaacacaa aacaac  |            |            |            |            |            |            |      |
| gttttcggat         gttcacqac         gccagtgtg         ttggaccqac         gcttacacac         cccctgttctg         gcttacacag         gcttacacac         cccagttctg         cgcttccgac         accacagacc         2100           acgacggaaat         gtttcacaac         gtttgggcta         ccacacttcac         cacattcacac         gaccgagaac         2220           tcaccacac         gcattgagac         acctcacacc         gccttcacacc         ccgtggttg         ccacctgacc         gcaccaacgc         tacttcacacc         ccatactacac         ccaccatcac         ccaccatcac         ccaccatcac         ccaggttg         tacgtcacac         catactgcc         2400           ccaccacatca         ccaggtttg         ccaggttcacc         cagtgcacc         atgtgatcac         ctgcacacgt         2400           ccaccatcac         ccaccactca         ccaccacgt         ctcaccaccac         cactactacac         cactactacac         cactactacac         cactactacac         cactactacac         cactacacac         cactacacacac         cactacacacac         cactacacacac         cactacacacacacacacacacacacacacacacacaca  |            |            |            |            |            |            |      |
| gtgttccggt         ttggaccga         gtttaggcta         cccaacgttg         cgtttccgac         acaaaagct         2100           acgcagcagtaa         gtttctggaa         aacgttacgg         gaaccatgg         2220           ttgaccagat         gcacgaagaa         atcatctccc         tgtgggaca         gtccttgaa         acagtgtgg         2220           caacgact         gcacgaagaa         atcatctccc         tgtgggaca         gtccttgaa         ccgtgcgtta         2340           ccaccatca         ccagacttg         ccgaaagtta         actctcaaa         ccatactgc         240           ccaccatca         ccaggettg         ccgaaagtta         actctcgac         gatcatcac         tgcaccatc         2520           gcacaaacgt         ttacgctatc         cagaaaaaga         aacacaaaac         ccttctcggt         aacaacgtta         2520           gcacacacgt         ttacgctatc         cagtgacaca         acaacatac         ccgttccgga         aacagttcg         2520           gcaaacagtta         accaccatt         ctgaaaaca         acaacatcat         ccgttccgaa         aacctpacca         2580           gtgtgcttga         agaaacatcaa         aacacacaa         aacacacaa         aacacacaa         aacaccacaa         26gttaccaa         gtgttacaa         gtftacacaat         <  |            |            |            |            |            |            |      |
| acgaccytgaacttcagacgtttcggatcccacgcttgcgttccgacgacccgaaca2160cgaagaaatcgttctggaaaacgttaccgaaaacttcaacattgtggaaaaacggaagac2220ttgaccagatgcacgaagacatccttcactgtgggaccagtccttgaaaccgtgtgtta2280accacacacaccaagactaaccagactaccacacacacccagacgatacgcaccaaccctaccttcaaaacacacgtta2340ccaccatacacccagcatcacccagacttacccagacttacccagacttacccagacttacccagacttaccgaaagttaacttcagcccgatcacagac2460ctccggttggttccaccgttcagtgacaccatggtacaccacttcaccgt2520gcaacacttaccagttccctggctgaaaaagaaatcacacccgttccgaaacacgttcgt2520taagagttaaacacatcaccgctgaaaaagaaatcacacccgttccgaaaccgttcgt2760ctaaggcttactgcaacatctccggatcaccaaaaactgctcgtgttacaacgtgttacaacgtgttacaacgtgttacaacgtgttacaacgtgttacaa2760ctaagacgttaatcaccacacaacacacacaaaacaccacaaaacctgaaacdtcaccctgccetcggg2820ttgaacacgtgttcaacacaaacacacacaaaacacacacaaaacctgacaccetcaccetg2940tcaaacacgtttcaacacaaacctgacacacggttsgtggttgatcacctgccggttsgtggg3060ctggtacactttcaccacacatccaacacacacggttsgtgggtaccacacacagccggttsgtgg3120gttcatcagactactgaacacagctgacacacagctgacac  |            |            |            |            |            |            |      |
| eggagagaategttetggaaaacgttacegaaagattacegaatgtgagaaaaacgacatgg2220ttgaccagatgcacgaagacateatetecetgtgggaccagtecettaaaaaccgtgette2280ccaacgacatgaacaaagaaatecettacegeaccaacgetaccettaaaaccaacgtta2340ccaacgacatgaacaaagaaatecettacegeteettegatgcatacatgcaacgtt2400ccaccatcacccaggettgccgaaagttaacaacaacaccteteceggtaaaggteceg2520gcaacaacgtttecaccgtcegtaaatgaacaacaaacacteteceggtaaaggteceg2580tgetgetgaaacgttecetggetgaaaaagaaaateatacacegttecegaaaaccttecegg2760ctaaggettactgcaacateteceggttecaacaaateacacegttecegaaaaccttacagg2760ctaagactgtaaacaacateacctcacctgaaacaactacacaacaacateacaacaactacaaaccttacaaacagg2760ctaagactggaaaacacacacacaacaaacaacaacateacaacaacateacaaccttacaaateacagacactacacateagacgaateaactacacateagacgaateaactacacateaactacacatea <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>   |            |            |            |            |            |            |      |
| ttgaccagat gcacgaagac atcatctocc tgtgggacca gtccttgaaa ccgtgcgtta 2380 aactgaccac gcbttgcgtt accttgact gcaccaacgc taccttcaaa aacaacgtta 2340 ccacaacgacat gaacaaagaa atccgtaact gcacctactac tgcaacgctt 2400 ccaccatcac ccaggcttg ccgaaagtta acttcgaccc gatccegate cactactgcg 2460 ctccggctgg ttacgctac ctgaaatgca acaacaaaaa cttccagctgt taccacgtt tccaccgtt cagtgacacc atggatcaa accggttgt tccaccgt 2520 gcaacaacgt ttccaccgtt ggtgaaaaag aaatcatcat ccgttccgaa aacctgacag 2540 acaacgttaa accatcatc gtgaaaaaga aaatcatcat ccgttccgaa aacctgacag 2640 acaacgttaa aaccatcatc ctgcaacaac accaccaaca attggacga accetgaaga 2760 ctagcactga aaccaccaca cactcctaa aaccatcaa attggacga agtgaccag 2880 ccacccgtt gtcaacaac cactccttca actggacgaa atcaccacac ggaaaactac aacacaaca aaccatcaa attggacgc 2640 acaacatga gaaaactac aacacaaca aaccatcaa attggacgg 2620 gtgacactgg aaccacacac acctccttca actggacggt gtgaattctt tacctggatg 2820 ctgtgacctgg aaccacggaa atcaccacc acctccttca actgccgtgg tgaattcttc tactgcacaca 2880 ccacccgtct gtcaacaaca caccacacac acctgacaga 2620 ctgtgcgtg 2620 ctgtgacaca 2620 ctgtgacgaa 2620 ctgggatggg 2620 ctgtgcgcg 2620 ctgtgcgaa 2620 ctgggtgggg 2620 ctgtgcgcg 2620 ctggggtggg 2620 ctgtgcgcg 2620 ctggggtggg 2620 ctggggtggg 2620 ctgggggggggggggggggggggggggggggggggggg   |            |            |            |            |            |            |      |
| aactgaaccc getgtgegt accetgeact geaccaacge tacetteaa acaacagtta 2340 ccaacgacat gaacaaagaa atcegtaact getecttege atgeatacaac tgeaacgett 2400 ccaccatcac caggettge cegaaagtta actetgacc gatecagate tacetactge 2460 cteeggetg ttacgette ctgaaatgca acaacaaaac ctteteeggt aaaggteeg 2520 gcaacaacgt ttecaacggt cagtgaacca atggtacaa accggttgtt tecaaccag 2580 tgetgetgaa eggtteect getgaaaag aaatcatcat cegtteegaa aaccatcage 2700 ctagegetta ctgcaacact teeggtteaa aaccatcaaca teeggttgt tecaaccagt 2700 ctagegetta ctgcaacaca teeggtteaa aaccatcaaca cacacacaa aaccatcaa attegeteeg 2700 ctagegettg aatcacacac cactcettea actagegetgg gaatactte teetggaaaactgg 2700 gtgacctgga aatcaccacc cactcettea actgcacgg gaaaactac cacacacaca aaccatcaa attegeteeg cegtgacagac 2880 ccaccegtet gtcaacaaca aaccgtaceg aagacgaaac catcacacat caccaccac cactcettea actgcacgg gaatactte teetggaaca 2880 ccaccegtet gtcaacaaca tecaacacac ceggtege tagtatette tactgcaaca 2880 ccaccagtac gacacacaa accgaagaa atctteegte gggtegtge tatgtacget ggetggtgg 3060 ctggtacaca aaccgaagaa atctteegte gggttggtg tacacatgaa gacaactggc 3120 gttcegaac gacaacaa aaccgaagaa atctteegte gggttggtg tacacatgaa gacaactggc 3120 gttcegaac gacagteeg getggtcage getggteeg ggatggtggtg ggaagaactaca accgacagaa accgttagaa gctcagtcaga gctcgtcaga gctcgtcage gctggtege ggtggtegg gacacgttega 3240 aacagctgca gacagttega gctcagtcaga gctcgaacaca aaccgaagaa atctecaca ceggtcagac gctgaacacaca aaccgaacaca gctcagaaca acagacagaca catgacagaca gcagaacaga gacagtggaacacaga acagagagaacacaga catgaacaga gacagaacaga   |            |            |            |            |            |            |      |
| ccacagacat gaacaaagaa atcegtaact getecttege atgeateaac tgeaacgett 2400 ccaccateac ccaggettge cegaaagta actetegacee gatecegate cactactgeg 2460 gcacacacgt ttecaccagt cagatagea acacaaaaac cttecteggt aaaggeteget 2520 gcaacaacgt ttecaccggt cagtgacace atggtateaa accggtegtt tecacccage 2580 tgetgetgaa cggttecetg getgaaaaag aaateateat ccgttecgaa aaccetgaceg 2700 ctagegetta ctgcaacate tecggtteca aaagaacga aaccetgaaa cggtgtataa 2760 aaaaactgaa ggaaaactaa aacaacaaca aaacaacaaa attegacege tecteceggtg 2820 gtgacetga aateaccace cactecttea acggetgg tgaattette tactgcaaca 2880 ccaccgtet gtteaacaac cactecttea acggetgg tgaattette tactgcaaca 2880 ccaccagtet gtteaacaac cactecttea acggetgg tgaattette tactgcaaca 280 ccaccagtet gtteaacaac acgetaceg aagagagaac cateacectg ccgtgcegta 2940 ctagaacaa aaccaaagaa atcttecgte ggggtggg taacatgaaa gacgagagggggtgggtggtcagggtgggtggtggtggggtgggggggg  |            |            |            |            |            |            |      |
| ccaccatcac ccaggettge ccgaaagta actecgace gatecegate cactactege 2460 ctecggetgg ttacgetate cdgaaatgca acaacaaaac ctecteggt aaaggteegt 2520 tgetgetgaa eggtteectg getgaaaaag aaateateat cegtteegaa aacetgace 2640 acaacgttaa accatcate ctecggtteea aaceateate teeggtteea aaceateate teeggtteea aaceateate teeggtteea aaceateate teeggtteea aateagaagega accetggaa eggtacetgg 2760 ctagecetga aaceateate teeggtteea aateagaagega accetggaa eggtacetgg 2760 aaaacetga ggaaaatea teeggtteea aateagaagega accetegtgg 2760 aaaacacga ggaaaacaca cactacteatea accaccaca aaceateaa aaceateaa gtgacetgga gaaaatea cactacacaca aaceateaa aaceateaa gtgacetgga aaceacaca cacteettea actgeetggg tgaatteette tactgeaaaa 2880 ccaccegtet gtteaacaac bggeaggggg ttggteggg tagtateette tactgeaaaa 2880 ccaccacgtet gtteaacaaca teeaacatea ceggtetget tagtgeetge ceggeggtgg 3060 aagacaacaa aaceagaagaa atctteegte gggttggtgg taacaatgaaga gacaaactgg 3120 gtteegaace gacaaaaaa aacegaagaa atctteegte eggttgetge getggtteet gegeggtggg 3060 aacagetgaa gacaactgg 262460 acctgetgg tgetateega geteagaaca catgetgea getggteete getggteete ggtggaaaaa aacegaagaa aaceggtaaaa catgacetgg 3180 aacagetgaa gacacetggt teggetategg acctgaaaa aacegaagaa aacegtgaaaa aacegaagaa aaceggtaaaa catgacetgg 3240 acctggaaaaa aacegaagaa aaceggaaaaaaaaaaaaaaa  |            |            |            |            |            |            |      |
| ctccggctgg ttacgctatc ctgaaatgca acaacaaac ctcccggt aaagtccg 2520 gcaacaacgt ttccaccgtt cagtgcaccc atggtatcaa accggttgtt tccacccage 2580 acaacacgttaa aaccatcatc gttcacctga aaaccatcaa ccgttcccga aaccatcaaccaccaccaccaccaccaccaccaccaccac   |            |            |            |            |            |            |      |
| gcaacaacgt tecaccgt cagtgcacc atggtatca accggttgt tecaccac 2580 tgctgctgaa acgttacca gttcaccacacacacacacacacacacacacacacacaca  |            |            |            |            |            |            |      |
| tgctgctgaa cggttcoctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2640 acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcgtt tgcacccgtg 2700 ctagcgctta ctgcaacatc tccggttcca aatggaacga aaccctgaaa cgtgttaaag 2760 aaaaactgca ggaaaactac aacaacaca aaaccatcaa attcgctccg tcctccggtg 2820 gtgacctgga aatcaccacc cactccttca actgccgtgg tgaattcttc tactgcaaca 2880 ccacccgtct gttcaacaac tccacaccac aaaccatcaa attcgctccg ccgtcgcgt 2940 tcaaacagat catcaacata tggcagggtg ttggtcgtc tatgtacgct ccgccgatcg 3000 ctggtacaca aaccgaagaa atcttccgtc cgggtggtgg tacaaacaca aaccgaagaa atcttccgtc gggtggtgg ggtggtgg ggcggtggtg gggggggg  |            |            |            |            |            |            |      |
| ctagogotta ctgcaacatc tccggttcca acagacga acaactcgt tgcaaccgtg 2700 ctagogotta ctgcaacatc tccggttcca aatggaacga accctgaaa cgtgttaaag 2760 acaacatcga ggaaacatc acacacaca aaccatca attcgctccg tcctccggtg 2820 ccacccgtct gttcaacac cactcatca actgccgtgg tgaattcttc tactgcacac 2880 ccacccgtct gttcaacaca caccaccac cactcatcac cactcacctg ccgtcgcta 2940 tcaacacagat caccaccag tggcaggggg ttggtgggggggggg  |            |            |            |            |            |            |      |
| ctagogotta ctgcaacatc tccggttcca aatggaacga aaccctgaaa cgtgttaaag 2760 aaaaactgca ggaaaactac aacaacaaca aaccatcaa attcgctccg tcctccggtg 2820 gtgacctgga aatcaccacc cactccttca acgcagtgg tgaattcttc tactgcaaca 2880 ccacccgtct gtcaacaac aacgctaccg aagacgaacc catcaccctg ccgtgccgta 2940 tcaaacaagat catcaacatg tggcagggg ttggtcgt tatgtacgct ccgccgatcg 3000 ctggtaacat caccagacaa tccaacatca ccggtctgct gctggttcgt gacggtggg 3060 aagacaacaa aaccgaagaa atcttccgtc cgggtgggg taacaatgaaa gacaactgg 3120 gttccgaact gtacaacatac aaagttatcg acatggaacc gctgggtatc gctccgaccg 3180 gtgctaaggt taccgatcag gctcgtcagc tacgtgctcc catcgtcag cagcagtcca 3240 aacagctgcg tgctatcgaa gctcagcagc acatgctgca gctggggggtaccggtgggggggggg   |            |            |            |            |            |            |      |
| aaaaactgca ggaaaactac aacaacaca aaaccatcaa attecetecg tectecggtg ggactetgga aatcaccacc cactecttca actgccgtgg tgaattette tactgcaaca 2880 ccaccegtet gttcaacaaca aacgctaeceg aagacgaac catcaccetg cegtgcegta 2940 tcaaacaagat catcaacatg tggacggtgg ttggtcgtgc tatgtacget cegcgateg 3000 ctggtaacat caccatgaa tcacacatg tggacggtgg ttggtcgtgt getggttegg gacggtggtg 3060 aagacaacaa aaccgaagaa atcttccgtc cgggtggtgg taacatgaaa gacaactggc 3120 gttccgaacet gtacaaatac aaagttatcg gctcgtcccc catcgtcaag gctcagtcagg tgctgtcccc gctgggtatc gctcgaccg 3180 accatgctgcg tgctatcaag gctcagcagc acctgctgcg tgctatcaga gctcagcagc acctgctgcg tgctgaccgt tctggaacaa catgactgg accggttt ctggctatcg acctgctgaaac catgaccgg accggttt ctggctatcg acctgtacag acctgctgaaac catgaccag accggtgaaac catgacaga accggaacaa catgaccag accggaacaa catgacaga accggaacaa catgacaga accggaacaa catgacaga accggaacaa accgaacaaa ctgtgggccc tctagaacaa ctgtgggccc tctagaacaa ctgtgggccc tctagaacaa ctgtgggccc agcaggaaaa aaacgaaaaa gacctgctgg 3480 accggaacaa catgacaga ccggaacaacaa ctgtgggccc tctagaacaa ctgtgggccc agcaccaca agcaggaaaa aaacgaaaaa gacctgctgg 3480 accggaacaa ctgtggaagac catggggaaca ctccaactgg ggaccaccaa ctgtgtgccc agcagaacaacaa ctgtgggggggggg   |            |            |            |            |            |            |      |
| gtgacctgga aatcacacc cactecttca actgccgtgg tgaattette tactgcaaca 2880 ccacccgtet gttcaacaac aacgctaccg aagacgaacc catcacctg ccgtgccgta 2940 tcaaacagat catcacactg tggcaggtg ttggtcgtgc tatgtacgct ccgccgatcg 3000 ctggtaacat cacctgcaaa tccaacaca ccggtctgct gctggttcgt gacggtgggg 3060 aagacaacaa aaccgaagaa atcttccgtc cgggtggg taacatgaaa gacaactgac 3120 gttccgaacc gtacaaatac aaagttatcg aactgaaaacc gctgggtatc gctccgaccg 3180 gtgctaagct taccgtcag gctcgtcagc tgctgtccc catcgttcag cagcagtcca 3240 acctgctgcg tgctatcgaa gctcagcagc acatgctgca gctgacagc tcgggaacc cagcagcagc aacagcagca cagcagcagc aacagcagca cagcagcagc acagcagcaaca cacgaagaac ctgggaaccc ctgggaaca ctccaactac accgaacaca 3420 acctggaccc ctgggaaca ctccaactac accgaacaca 3420 acctggaccc ctggaaaac ctgggaacacc ctggaaaaa aaacgaaaaa gacctgctgg 3480 gatccagacc ctggaaaaac ctgggaacac ctggaagaac ctgggaacaccaca 3420 agcaccagaac ccggaacac ctggaacacca 3420 agcaccagaac ctggaagac ctggaaaac ctgggaaaac ctggaagaac ctggaagac ccggaaaaa aaacgaaaaa gacctgctgg 3480 gatccagacc ctggaagag tgcaactcca agcacacaca 3420 agcaccagacc ctggaagag tgcaactcca agcaccaca accgacacac accgacacca accgacacacc accacacaca   | ctagcgctta | ctgcaacatc | tccggttcca | aatggaacga | aaccctgaaa | cgtgttaaag | 2760 |
| ccacccgtct gttcaacaac aacgetaccg aagacgaaac catcacctg ccgtgccgta 2940 tcaaacagat catcaacatg tggcagggtg ttggtcgtgc tatgtacgct ccgccgatcg 3000 ctggtaacat cacctgcaaa tccaacatca ccggtctgct gctggttcgt gacggtggtg 3060 aagacaacaa aaccgaagaa atcttccgtc cgggtggtgg taacaatgaa gacaactggc 3120 gttccgaacc gtacaaatac aaagttatcg aactgacacc gctgggtatc gctccgaccg 3180 gtggtaagct taccggtcag gctgtacgc tgctgtcgc catcgttcag cagcagctga gctgatcag gctgatcag gctgatcag gctgacagca cagctgctgg gacagctgca gaccgtgtt ctggctatcg aacgttacct gaaagaccag cagctgaga 3360 tctggaacaa catgacctgg atggaatgg accgtgaaa catgaccag acatgctgaa accgtgaaca catgacctgg atggaatgg accgtgaaaa catgaccag acggaagaaa aaacgaaaaa gacctgctgg 3480 ctctggaaca ctggaagac tcccagaccc agcaggaaaaa catgaccaga ctggggtcc ggttcgacat ctccaactac accgacacca 3420 tctaccgtct gctggaagac tccagaccc agcaggaaaaa aaacgaaaaa gacctgctgg 3480 ctctggaact tgctgtgct tctagttgcc agcactctgt tgtttgccc tcacactag gaggaaattg aggaggagga tgggagggg tggcactcca ctgtccttc ctaataaaaa gaggaaattg 3660 ctcctctgac cctggaagg gcacatccca agccactcgt tctgggtggg tgggggggg caggaaattg aggaggagga ttgggaagga aatagaggc aggacagga agaaattga cccggttcct cctggggcag aaagaagagg gcacatccca 3840 tctcttgtga cacaccctgt ccaggtcct cctggggcag aaagaagagg gcacatccca 3840 tctcttgtga cacaccctgt ccaggttcct cctggggcag aaagaagagg gcacatccca 3840 tctcttgtga cacaccctgt ccaggttcct cctggggcag aaagaagagg gcacatccca 3840 tctctctgtga cacaccctgt ccaggttcct cctggggcag aaagaagagg gcacatccca 3840 tctctctgtga cacaccctgt ccaggttcct cctgggccag aaagaagagg gcacatcca 3840 tctcctcccc catcagcca caaaccaaa cctagcccc aaagaagaga gaaaattaaa 4020 gcaagaaaaagg ctattaaagg caaaccaaa cctaggcca tcaaggcagga agaaaattaaa 4020 gcaagaaaaagg ctattaaggt tatccacaga ttaaggcca tcaagggga aacacacggg aaaaaggc cagaaaagg cagaaaagg cagaaacgg cagaaaagg cagaaacgg tatcaagga aaagggaga aacaacggga aacaacggga agaaactgg  tatcagcaa aaggaggaga aacaacggga aacaaggga agaaaatgg  cagaaaagg  cagaaaagg  cagaaacag  caggagaacg  caggaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  | aaaaactgca | ggaaaactac | aacaacaaca | aaaccatcaa | attcgctccg | tcctccggtg | 2820 |
| ccacccgtct gttcaacaac aacgetaccg aagacgaaac catcacctg ccgtgccgta 2940 tcaaacagat catcaacatg tggcagggtg ttggtcgtgc tatgtacgct ccgccgatcg 3000 ctggtaacat cacctgcaaa tccaacatca ccggtctgct gctggttcgt gacggtggtg 3060 aagacaacaa aaccgaagaa atcttccgtc cgggtggtgg taacaatgaa gacaactggc 3120 gttccgaacc gtacaaatac aaagttatcg aactgacacc gctgggtatc gctccgaccg 3180 gtggtaagct taccggtcag gctgtacgc tgctgtcgc catcgttcag cagcagctga gctgatcag gctgatcag gctgatcag gctgacagca cagctgctgg gacagctgca gaccgtgtt ctggctatcg aacgttacct gaaagaccag cagctgaga 3360 tctggaacaa catgacctgg atggaatgg accgtgaaa catgaccag acatgctgaa accgtgaaca catgacctgg atggaatgg accgtgaaaa catgaccag acggaagaaa aaacgaaaaa gacctgctgg 3480 ctctggaaca ctggaagac tcccagaccc agcaggaaaaa catgaccaga ctggggtcc ggttcgacat ctccaactac accgacacca 3420 tctaccgtct gctggaagac tccagaccc agcaggaaaaa aaacgaaaaa gacctgctgg 3480 ctctggaact tgctgtgct tctagttgcc agcactctgt tgtttgccc tcacactag gaggaaattg aggaggagga tgggagggg tggcactcca ctgtccttc ctaataaaaa gaggaaattg 3660 ctcctctgac cctggaagg gcacatccca agccactcgt tctgggtggg tgggggggg caggaaattg aggaggagga ttgggaagga aatagaggc aggacagga agaaattga cccggttcct cctggggcag aaagaagagg gcacatccca 3840 tctcttgtga cacaccctgt ccaggtcct cctggggcag aaagaagagg gcacatccca 3840 tctcttgtga cacaccctgt ccaggttcct cctggggcag aaagaagagg gcacatccca 3840 tctcttgtga cacaccctgt ccaggttcct cctggggcag aaagaagagg gcacatccca 3840 tctctctgtga cacaccctgt ccaggttcct cctggggcag aaagaagagg gcacatccca 3840 tctctctgtga cacaccctgt ccaggttcct cctgggccag aaagaagagg gcacatcca 3840 tctcctcccc catcagcca caaaccaaa cctagcccc aaagaagaga gaaaattaaa 4020 gcaagaaaaagg ctattaaagg caaaccaaa cctaggcca tcaaggcagga agaaaattaaa 4020 gcaagaaaaagg ctattaaggt tatccacaga ttaaggcca tcaagggga aacacacggg aaaaaggc cagaaaagg cagaaaagg cagaaacgg cagaaaagg cagaaacgg tatcaagga aaagggaga aacaacggga aacaacggga agaaactgg  tatcagcaa aaggaggaga aacaacggga aacaaggga agaaaatgg  cagaaaagg  cagaaaagg  cagaaacag  caggagaacg  caggaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  cagaaaagg  | gtgacctgga | aatcaccacc | cactccttca | actgccgtgg | tgaattcttc | tactgcaaca | 2880 |
| tcaaacagatcatcaacatgtggcaggtgttggtcgtgtatgtacgetccgccgateg3000ctggtaacatcacctgcaaatccaacatcaccggttggtggacggttggtg3060aagacaacaaaaccagaagaaatcttccgtccgggtggtggtaacatgaaagacaactggc3120gtccaacatgtacaaatacaacgtacagctgctgtcaccctgggtatcgctccgaccg3180gtgtaagettaccgttcaggctcgtcagccactgtcagccactgtcagcactgtcagccagcagtcca3240aacagctgcagacccgtgttctggctatcgacatgctgcagctgaccgtttggggtatca3300acaggtgaacacatgacctggatggaatggaccgtgaaatctccaactacaccgacacca3420tctaccgtctgctggaagacctccaagaccagcaggaaaaaaacgaaaaagacctgctgg3540gatccagatcctggaagagctgtggtcctggttcgacatctcaaactagctgtggtggg3540gatccagatctctgagaagggccactcccactgtctttcctaataaaaagacggaaattg3660cttctcttgaccttgagtaggtgtcattctatctctgggggtcggtggggcaggacagca3770agggggagagatttgggaagacaatagcaccaatgctgggggtctattggga3360ctccttgtgacagaagattgacccggttcctcttgggggggctattaggac3780ctcattgggagagatttgggaagacatgctgggggtctattgggggctattaggac3780ctcattgggtcgaagaattgacccggttcctcttgggggggctataggac3780ctcattggctcagaagaggaccccggttctcactt  |            |            |            |            |            |            |      |
| ctggtaacat cacctgcaaa tccaacatca ccggtctgct gctggttcgt gacggtggtg 3060 aagacaacaa aaccgaagaa atcttccgtc cgggtggtgg taacatgaaa gacaactggc 3120 gttccgaact gtacaaatac aaagttatcg aactgaaacc gctgggtatc gctccgaccg 3180 gtgctaagct taccgttcag gctcgtcagc tgctgtcctc catcgttcag cagcagtcca 3240 acctgctgcg tgctatcgaa gctcagcag aactgctgca gctgaccgtt tggggtatca 3300 aacagctgca gacccgtgtt ctggctatcg aacgttacct gaaagaccag cagctcgaga 3360 tctggaacaa catgacctgg atggaatgg accgtgaaaa ctccaactac accgacacc 3420 ctctggaacac ctggaaagac tcccaagaccc agcaggaaaa aaacgaaaaa gacctgctgg 3480 gatccagatc ctggaaaaac ctgtggtcct ggttcgacat ctccaactac accgacacca 3420 gatccagatc tgctggcct tctagttgcc agcaatctgt tgtttgccc tcccactgg 3560 catcgcattg tctggaagac gccactcca ctgtcgtctc ctaataaaat gaggaaattg 3660 catcgcattg tctgagtagg tgtcattcta tctgggggg tgggggggg caggacagca 3720 agggggagga ttgggaagac aatagcaggc atgctgggg tgggggggggg  |            |            |            |            |            |            |      |
| aagacaacaa aaccgaagaa atcttccgtc cgggtggtgg taacatgaaa gacaactggc 3120 gttccgaact gtacaaatac aaagttatcg aactgaaacc gctgggtatc gctccgaccg 3180 gtgctaagct taccgttcag gctcgtcagc tgctgtcctc catcgttcag cagcagtcca 3240 acctgctgcg tgctatcgaa gctcagcag acatgctgca gctgaccgtt tggggtatca 3300 accagctgca gacccgtgtt ctggaacac catgacacca accgacacca catgacctgg atggaatggg accgtacct gaaagaccag cagctcgaga 3360 tctaccgtct gctggaagac tcccaagaccc agcaggaaaa accgacacaca accgacacca 3420 tctaccgtct gctggaagac tcccaagaccc agcaggaaaaa cctgcaactac ctcgaactcc tgggaagac tcccaggaccc agcaggaaaaa aaccgaaaaaa gacctgctgg 3480 gatccagatc tgctgggct tctagtggccc tctagtggccc tctagtggccc tctagtggccc tctagtggccc tctagtggccc tctagtggcc tctagtggcc tctagtgggggggggg  |            |            |            |            |            |            |      |
| gttccgaact gtacaaatac aaagttatcg gctcgtacgc tgctgtcctc catcgttcag cagcagtcca 3240 acctgctgcg tgctatcgaa gctcagcag acatgctgca gctgaccgt tggggtatc gccgagtcca 3240 acctgctgcg tgctatcgaa gctcagcag acatgctgca gctgaccgtt tggggtatca 3300 accagctgca gacccgtgtt ctggctatcg acctgtacct gaaagaccag cagctcgaga 3360 tctggaacaa catgacctgg atggaatggg accgtgaaaa ctccaactac accgacacca 3420 tctaccgtct gctggaagac tcccagaccc agcaggaaaa aaacgaaaaa gacctgctgg 3480 ctctggactc ctggaaaaac ctgtggtcct ggttcgacat ctccaactgg ctgtggtgag 3540 gatccagatc tgctgtgcct tctagttgcc agccatctgt tgtttgccc tccccegtgc 3600 cttccttgac cctggaaggt gccactccaa ctgtgcttc ctaaaaat gaggaaaattg 3660 catcgcattg tctgagtagg tgtcattcta ttctgggggg tgggggggggg  |            |            |            |            |            |            |      |
| gtgctaagct taccgttcag gctcgtcagc tgctgtcctc catcgttcag cagcagtcca 3240 acctgctgcg tgctatcgaa gctcagcagc acatgctgca gctgaccgtt tggggtatca 3300 aacagctgca gacccgtgtt ctggctatcg accgttacct gacagcagc cagctcgaga 3360 tctggaacaa catgacctgg atggaatggg accgtgaaat ctccaactac accgacacca 3420 tctaccgtct gctggaagac tcccagaccc agcaggaaaa aaacgaaaaa gacctgctgg 3480 ctctggactc ctggaagac ctgtggtcct ggttcgacat ctccaactgg ctgtggtgag 3540 gatccagatc tgctgtgcct tctagttgcc agccatctgt tgtttgccc tcccccgtgc 3600 cttccttgac cctggaaggt gccactccca ctgtccttc ctaataaaat gaggaaattg 3660 catcgcattg tctgagtagg tgtcatcta ttctgggggg tggggtgggg  |            |            |            |            |            |            |      |
| acctgctgcg tgctatcgaa gctcagcagc acatgctgca gctgaccgtt tggggtatca 3300 aacagctgca gacccgtgtt ctggctatcg aacgttacct gaaagaccag cagctcgaga 3360 tctggaacaa catgacctgg atggaatggg accgtgaaat ctccaactac accgacacca 3420 tctaccgtct gctggaagac tcccagaccc agcaggaaaa aaacgaaaaa gacctgctgg 3480 ctctggactc ctggaaaaac ctgtggtcct ggttcgacat ctccaactag ctgtggtgag 3540 gatccagatc tgctgtgcct tctagttgcc agccatctgt tgtttgcccc tccccegtgc 3600 cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg 3660 catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg  |            |            |            |            |            |            |      |
| aacagctgca gacccgtgtt ctggctatcg aacgttacct gaaagaccag cagctegaga 3360 tctggaacaa catgacctgg atggaatggg accgtgaaat ctccaactac accgacacca 3420 tctaccgtct gctggaagac tcccagaccc agcaggaaaa aaacgaaaaa gacctgctgg 3480 ctctggactc ctggaaaaac ctgtggtcct ggttcgacat ctccaactgg ctgtggtgag 3540 gatccagatc tgctgtgcct tctagttgcc agccatctgt tgtttgecec tccccegtgc 3600 cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg 3660 catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg   |            |            |            |            |            |            |      |
| tctggaacaa catgacetgg atggaatggg accgtgaaat ctccaactac accgacacca 3420 tctaccgtct gctggaagac tcccagaccc agcaggaaaa aaacgaaaaa gacctgctgg 3480 ctctggactc ctggaaaaac ctgtggtcct ggttcgacat ctccaactgg ctgtggtgag 3540 gatccagatc tgctgtgcct tctagttgcc agccatctgt tgtttgcccc tccccegtgc 3600 cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaaattg 3660 catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg   |            |            |            |            |            |            |      |
| tetacegtet getggaagac teccagacec agcaggaaaa aaacgaaaaa gacetgetgg 3480 ctetggacte etggaaaaac etgtggteet ggttegacat etccaactgg etgtggtgag 3540 gatecagate tgetgtgeet tetagttgee agceatetgt tgtttgeee tecceegtge 3600 etteettgae eetggaaggt gecacteea etgteettte etaataaaat gaggaaattg 3660 categeattg tetggatgag tgteatteta ttetgggggg tggggtgggg   |            |            |            |            |            |            |      |
| ctetggacte ctggaaaaac ctgtggtcet ggttcgacat ctccaactgg ctgtggtgag 3540 gatccagate tgctgtgcct tctagttgce agccatctgt tgtttgccc tcccccgtgc 3600 cttccttgac cctggaaggt gccactcca ctgtcctttc ctaataaaat gaggaaattg 3660 catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg  |            |            |            |            |            |            |      |
| gatecagate tgetgtgeet tetagttgee agecatetgt tgtttgeee teeecegtge 3600 etteettgae cetggaaggt gecacteea etgteettte etaataaaat gaggaaattg 3660 eategeattg tetgagtagg tgteatteta ttetgggggg tggggtgggg   |            |            |            |            |            |            |      |
| ctteettgae cetggaaggt gecaetecea etgteettte etaataaaat gaggaaattg 3660 categeattg tetgagtagg tgteatteta ttetgggggg tggggtgggg  |            |            |            |            |            |            |      |
| categoatty tetgagtagy tyteatteta ttetgggggg tygggtgggg caggacagca 3720 agggggagga ttgggaagac aatagcaggc atgetggga tygggtgggg tetatgggta 3780 cecaggtget gaagaattga eceggteet cetgggecag aaagaagcag geacatecec 3840 ttetetytga cacaccetyt ceaegeceet gyteetagt teeageceea eteataggac 3900 acteataget caggaggget eegeetteaa teecaeeege taaagtaett gyageggtet 3960 eteeeteet cateageca ecaaaccaaa eetageetee aaggatggga agaaattaaa 4020 geaagatagg etattaagtg cagagggag gaaaatgeet eeaaggegga agaaattaaa 4080 gagagaaate atagaattt aaggecatga tetaaggeea teatggeet aateeteege 4140 teeetegete aetgaetege tygggteggga agaaatggaga agaacatgtg 4260 ageaaaagge gaaaagg eeaggaaceg taaeaggeggaaaagge eegeaaaagg eeaggaaceg taaaaaggee teaaggegaa ageateeeg tyggeteege teaagteaga gytggegaaa 4380 eeegacagga etataaagat accaggegtt teeeeetgga ageteeeteg tyggeteee 4440  |            |            |            |            |            |            |      |
| agggggaga ttgggaagac aatagcaggc atgctggga tgcggtgggc tctatgggta 3780 cccaggtgct gaagaattga cccggttcct cctgggccag aaagaagcag gcacatcccc 3840 ttctctgtga cacaccctgt ccacgccct ggttcttagt tccagccca ctcataggac 3900 actcatagct caggagggct ccgccttcaa tcccacccgc taaagtactt ggagcggtct 3960 ctccctcct catcagccca ccaaaccaaa cctagcctc aaggatggga agaaattaaa 4020 gcaagatagg ctattaagtg cagagggag gaaaatgcct ccaacatgtg aggaagtaat 4080 gagagaaatc atagaattt aaggccatga tttaaggcca tcatggcct aatcttccgc 4140 ttcctcgctc actgactcgc tgcgctcggt cgttcggct cggcgaggga agaacatgtg 4260 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgttttcca 4320 taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 4380 cccgacagga ctataaagat accaggcgtt tcccctgga agctccctcg tgcgctcc 4440  |            |            |            |            |            |            |      |
| cccaggtgct gaagaattga cccggttcct cctgggccag aaagaagcag gcacatcccc 3840 ttctctgtga cacaccctgt ccacgccct ggttcttagt tccagccca ctcataggac 3900 actcatagct caggagggct ccgccttcaa tcccacccgc taaagtactt ggagcggtct 3960 ctccctcct catcagccca ccaaaccaaa cctagcctc aagagtggga agaaattaaa 4020 gcaagatagg ctattaagtg cagagggaga gaaaatgcct ccaacatgtg aggaagtaat 4080 gagagaaatc atagaatttt aaggccatga tttaaggcca tcatggcctt aatcttccgc 4140 ttcctcgctc actgactcgc tgcgctcggt cgttcggct cggcgagggg tatcagctca 4200 ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 4260 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgttttcca 4320 taggctcgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 4380 cccgacagga ctataaagat accaggcgtt tcccctgga agctcctcg tgcgctccc 4440  |            |            |            |            |            |            |      |
| ttetetgtga cacaccetgt ceaegeceet ggttettagt teeageceea eteataggae 3900 acteataget caggaggget cegeetteaa teecaecege taaagtaett ggageggtet 3960 eteecteet cateagecea ecaaaceaaa eetageetee aagagtggga agaaattaaa 4020 geaagataagg etattaagtg cagagggaga gaaaatgeet eeaaeatgtg aggaagtaat 4080 gagagaaate atagaatttt aaggeeatga titaaggeea teatggeett aateeteege 4140 eteetegete aetgaetege tgegeteggt egteegget eggegagegg tateagetea 4200 eteaaaggeg gtaataeggt tateeacaga ateaggggat aaegeaggaa agaacatgtg 4260 ageaaaagge eageaaaagg eeaggaaeeg taaaaaggee gegttgetgg egtttteea 4320 taggeteege eeeeetgae ageateacaa aaategaege teaagteaga ggtggegaaa 4380 eeegacagga etataaagat aceaggegtt teeeeetgga ageteeeteg tgegetetee 4440   |            |            |            |            |            |            |      |
| acteataget caggaggget cegeetteaa teceaceege taaagtaett ggageggtet 3960 eteeeteet cateageeca ecaaaceaaa cetageetee aagagtggga agaaattaaa 4020 gcaagatagg etattaagtg cagagggaga gaaaatgeet ecaacatgtg aggaagtaat 4080 gagagaaate atagaattt aaggeeatga tetaaggeea teatggeett aatetteege 4140 eteetegete actgaetege tgegeteggt egteegget eggegaggggaggaggaggaggaggaggaggaggaggagg  |            |            |            |            |            |            |      |
| ctcctcct catcagcca ccaaaccaaa cctagcctc aagagtgga agaaattaaa 4020 gcaagatagg ctattaagtg cagagggaga gaaaatgcct ccaacatgtg aggaagtaat 4080 gagagaaatc atagaattt aaggccatga bttaaggcca tcatggcctt aatcttccgc 4140 ttcctcgct actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca 4200 ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 4260 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgttttcca 4320 taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 4380 cccgacagga ctataaagat accaggcgtt tcccctgga agctccctcg tgcgctctcc 4440  |            |            |            |            |            |            |      |
| gcaagatagg ctattaagtg cagagggaga gaaaatgcct ccaacatgtg aggaagtaat 4080 gagagaaatc atagaattt aaggccatga bttaaggcca tcatggcctt aatcttccgc 4140 ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca 4200 ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 4260 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgttttcca 4320 taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 4380 cccgacagga ctataaagat accaggcgtt tcccctgga agctccctcg tgcgctctcc 4440   |            |            |            |            |            |            |      |
| gagagaaatc atagaattt aaggccatga ttaaggcca tcatggcct aatcttccgc 4140 ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca 4200 ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 4260 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgttttcca 4320 taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 4380 cccgacagga ctataaagat accaggcgtt tcccctgga agctccctcg tgcgctctcc 4440  |            |            |            |            |            |            |      |
| ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca 4200 ctcaaaggcg gtaatacggt tatcacaga atcaggggat aacgcaggaa agaacatgtg 4260 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 4320 taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 4380 cccgacagga ctataaagat accaggcgtt tcccctgga agctccctcg tgcgctctcc 4440  |            |            |            |            |            |            |      |
| ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 4260 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 4320 taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 4380 cccgacagga ctataaagat accaggcgtt tcccctgga agctccctcg tgcgctctcc 4440  |            |            |            |            |            |            |      |
| agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 4320 taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 4380 cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 4440  |            |            |            |            |            |            |      |
| taggeteege eeceetgaeg ageateacaa aaategaege teaagteaga ggtggegaaa 4380 eecgacagga etataaagat accaggegtt teeceetgga ageteeeteg tgegetetee 4440  |            |            |            |            |            |            |      |
| cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 4440   |            |            |            |            |            |            |      |
|  |            |            |            |            |            |            |      |
| tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 4500   |            |            |            |            |            |            |      |
|  | tgttccgacc | ctgccgctta | ccggatacct | gtccgccttt | ctcccttcgg | gaagcgtggc | 4500 |

```
gettteteat ageteacget gtaggtatet cagtteggtg taggtegtte getecaaget 4560
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 4620
tettgagtee aacceggtaa gacacgaett ategecaetg geageageea etggtaacag 4680
gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta 4740
cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 4800
aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 4860
tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt 4920
ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag 4980
attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat 5040
ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc 5100
tateteageg atetgtetat ttegtteate catagttgee tgactegggg ggggggggg 5160
ctgaggtctg cctcgtgaag aaggtgttgc tgactcatac caggcctgaa tcgccccatc 5220
atccagccag aaagtgaggg agccacggtt gatgagagct ttgttgtagg tggaccagtt 5280
ggtgattttg aacttttgct ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat 5340
ctgatectte aacteageaa aagttegatt tatteaacaa ageegeegte cegteaagte 5400
agogtaatgo totgocagtg ttadaaccaa ttaaccaatt otgattagaa aaactcatog 5460
agcatcaaat gaaactgcaa tttattcata tcaggattat caataccata tttttgaaaa 5520
agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc 5580
tggtatcggt ctgcgattcc gactcgtcca acatcaatac aacctattaa tttcccctcg 5640
tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat 5700
ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca 5760
tcaaaatcac tcgcatcaac caaaccgtta ttcattcgtg attgcgcctg agcgagacga 5820
aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa ccggcgcagg 5880
aacactgcca gegeateaac aatattttea eetgaateag gatattette taataeetgg 5940
aatgetgttt teeeggggat egeagtggtg agtaaceatg cateateagg agtaeggata 6000
aaatgettga tggteggaag aggeataaat teegteagee agtttagtet gaeeatetea 6060
tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc tggcgcatcg 6120
ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc gcgaqcccat 6180
ttatacccat ataaatcagc atccatgttg gaatttaatc gcggcctcqa qcaaqacqtt 6240
tecegitigaa tatggeteat aacacceett gtattaetgt ttatgtaage agacagtttt 6300
attgttcatg atgatatatt tttatcttgt gcaatgtaac atcaqagatt ttgaqacaca 6360
acgtggcttt ccccccccc ccattattga agcatttatc agggttattg tctcatgagc 6420
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 6480
cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat 6540
aggogtatea egaggeeett tegte
                                                                  6565
```

```
<211> 6526
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 137
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgg tcagcggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
```

ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480

<210> 137

| WU 02/3294      | <b>ა</b>   |            |             |             | PC1/US     | 01/25/2 |
|-----------------|------------|------------|-------------|-------------|------------|---------|
| catagtaacg      | ccaataggga | ctttccattg | acgtcaatgg  | gtggagtatt  | tacggtaaac | 540     |
| tgcccacttg      | gcagtacatc | aagtgtatca | tatgccaagt  | acgcccccta  | ttgacgtcaa | 600     |
| tgacggtaaa      | tggcccgcct | ggcattatgc | ccagtacatg  | accttatggg  | actttcctac | 660     |
| ttggcagtac      | atctacgtat | tagtcatcgc | tattaccatg  | gtgatgcggt  | tttggcagta | 720     |
| catcaatggg      | cgtggatagc | ggtttgactc | acggggattt  | ccaagtctcc  | accccattga | 780     |
| cgtcaatggg      | agtttgttt  | ggcaccaaaa | tcaacgggac  | tttccaaaat  | gtcgtaacaa | 840     |
| ctccgcccca      | ttgacgcaaa | tgggcggtag | gcgtgtacgg  | tgggaggtct  | atataagcag | 900     |
| agctcgttta      | gtgaaccgtc | agatcgcctg | gagacgccat  | ccacgctgtt  | ttgacctcca | 960     |
| tagaagacac      | cgggaccgat | ccagcctccg | cggccgggaa  | cggtgcattg  | gaacgcggat | 1020    |
| tccccgtgcc      | aagagtgacg | taagtaccgc | ctatagactc  | tataggcaca  | cccctttggc | 1080    |
| tcttatgcat      | gctatactgt | ttttggcttg | gggcctatac  | acccccgctt  | ccttatgcta | 1140    |
| taggtgatgg      | tatagcttag | cctataggtg | tgggttattg  | accattattg  | accactcccc | 1200    |
| tattggtgac      | gatactttcc | attactaatc | cataacatgg  | ctctttgcca  | caactatctc | 1260    |
| tattggctat      | atgccaatac | tctgtccttc | agagactgac  | acggactctg  | tatttttaca | 1320    |
| ggatggggtc      | ccatttatta | tttacaaatt | cacatataca  | acaacgccqt  | ccccgtgcc  | 1380    |
| cgcagttttt      | attaaacata | gcgtgggatc | tecaegegaa  | tctcgggtac  | gtgttccgga | 1440    |
| catgggctct      | tctccggtag | cggcggagct | tccacatccg  | agccctqqtc  | ccatgectee | 1500    |
| agcggctcat      | ggtcgctcgg | cageteettg | ctcctaacaq  | tggaggccag  | acttaggcac | 1560    |
| agcacaatgc      | ccaccaccac | cagtgtgccg | cacaaqqccq  | tagcagtagg  | gtatgtgtct | 1620    |
| gaaaatgagc      | gtggagattg | ggctcgcacg | gctgacgcag  | atggaagact  | taaggcagcg | 1680    |
| gcagaagaag      | atgcaggcag | ctgagttgtt | gtattctgat  | aagagtcaga  | ggtaactccc | 1740    |
| gttgcggtgc      | tgttaacggt | ggagggcagt | gtagtctgag  | cagtactcgt  | tgctgccgcg | 1800    |
| cgcgccacca      | gacataatag | ctgacagact | aacagactgt  | tcctttccat  | gggtcttttc | 1860    |
| tgcagtcacc      | gtcgtcgaca | cqtqtqatca | gatatogogg  | ccactctaga  | gatatcgcca | 1920    |
| ccatgcgtgt      | togtggtato | ccqcqtaact | ggccgcagtg  | gtggatgtgg  | ggtatcctgg | 1980    |
| gtttctggat      | gatcatcatc | taccatatta | ttggtaacat  | gtgggttacc  | gtttactacg | 2040    |
| gtgttccggt      | ttggaccgac | gctaaaacca | ccctattcta  | cacttccac   | accaaagcct | 2100    |
| acgaccgtga      | agttcacaac | gtttgggcta | cccacactta  | cattccaacc  | gacccgaacc | 2160    |
| cgcaggaaat      | cattetagaa | aacqttaccq | aaaacttcaa  | catotogaaa  | aacgacatgg | 2220    |
| ttgaccagat      | gcacqaaqac | atcatctccc | tataggacca  | gtccctgaaa  | ccgtgcgtta | 2280    |
| aactgacccc      | gctgtgcgtt | accetquact | gcaccaacgc  | taccttcaaa  | aacaacgtta | 2340    |
| ccaacgacat      | gaacaaagaa | atccqtaact | gctccttcgc  | atgcatcaac  | tgcaacgctt | 2400    |
| ccaccatcac      | ccaggettge | ccgaaagtta | acttcgaccc  | gatecegate  | cactactgcg | 2460    |
| ctccggctgg      | ttacqctatc | ctgaaatgca | acaacaaaac  | cttctccaat  | aaaggtccgt | 2520    |
| gcaacaacgt      | ttccaccgtt | cagtgcaccc | atggtatcaa  | acconttatt  | tccacccagc | 2580    |
| tgctgctgaa      | cggttccctq | gctgaaaaag | aaatcatcat  | ccqttccqaa  | aacctgaccg | 2640    |
| acaacgttaa      | aaccatcatc | gttcacctga | acaaatccqt  | tgaaatcgtt  | tgcacccgtg | 2700    |
| ctagcgctta      | ctgcaacatc | tccqqttcca | aatggaacga  | aaccctgaaa  | cgtgttaaag | 2760    |
| aaaaactgca      | ggaaaactac | aacaacaaca | aaaccatcaa  | attcactcca  | tcctccggtg | 2820    |
| gtgacctgga      | aatcaccacc | cactccttca | actoccotoo  | tgaattcttc  | tactgcaaca | 2880    |
| ccgctagcct      | gccgtgccgt | atcaaacaga | tcatcaacat  | gtagcagggt  | gttggtcgtg | 2940    |
| ctatgtacgc      | tecgeegate | gctggtaaca | tcacctgcaa  | atccaacatc  | accggtctgc | 3000    |
| tgctggttcg      | tgacggtggt | gaagacaaca | aaaccgaaga  | aatcttccgt  | ccgggtggtg | 3060    |
| gtaacatgaa      | agacaactgg | cqttccqaac | tgtacaaata  | caaagttatc  | gaactgaaac | 3120    |
|                 |            |            |             |             | ctgctgtcct |         |
| ccatcgttca      | gcagcagtcc | aacctgctgc | gtgctatcga  | ageteageag  | cacatgctgc | 3240    |
| agetgaceqt      | ttggggtatc | aaacagctgc | agacccgtgt  | tctggctatc  | gaacgttacc | 3300    |
| tgaaagacca      | gcagetegag | atctggaaca | acatgacctg  | gatggaatgg  | gaccgtgaaa | 3360    |
| tctccaacta      | caccgacacc | atctaccate | tactagaaga  | ctcccagacc  | cagcaggaaa | 3420    |
| aaaacgaaaa      | agacctacta | gctctqqact | cctggaaaaa  | cetataatee  | tggttcgaca | 3480    |
| tctccaacto      | gctgtgatga | ggatccacat | ctactatace  | ttctagttgc  | cagccatctg | 3540    |
| ttgtttaccc      | ctccccata  | ccttccttaa | ccctggaaagg | toccactece  | actgtccttt | 3600    |
| cctaataaaa      | tgaggaaatt | gcatcgcatt | atctgagtag  | atatcattct  | attctggggg | 3660    |
| gtaggataga      | gcaggacagc | aaqqqqqaaq | attoggaaga  | caatagcagg  | catgctgggg | 3720    |
| atgcqqtqqq      | ctctataaat | acccadatac | tgaagaattg  | accognition | tcctgggcca | 3720    |
| - 5 - 55 - 55 5 |            |            | - Junganicy | 2000990000  | cccigggcca | טפוב    |

```
gaaagaagca ggcacateee ettetetgtg acacaceetg tecaegeeee tggttettaq 3840
ttccagcccc actcatagga cactcatage tcaggaggge tccgccttca atcccacccg 3900
caagagtggg aagaaattaa agcaagatag gctattaagt gcagagggag agaaaatgcc 4020
tccaacatgt gaggaagtaa tgagagaaat catagaattt taaggccatg atttaaggcc 4080
atcatggcct taatcttccg cttcctcgct cactgactcg ctgcqctcqq tcqttcqqct 4140
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 4200
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 4260
egegttgetg gegtttttee ataggeteeg ceceetqae qaqeateaca aaaateqaeq 4320
ctcaaqtcaq aqqtqqcqaa acccqacaqq actataaaqa taccaqqcqt ttcccctqq 4380
aageteeete gtgegetete etgtteegae eetgeegett aeeggataee tgteegeett 4440
tetecetteg ggaagegtgg egetttetea tageteaege tgtaggtate teagtteggt 4500
gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccqctg 4560
egecttatee ggtaactate gtettgagte caaceeggta agacaegact tategecact 4620
ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 4680
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta tctgcgctct 4740
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 4800
cgctqqtaqc qqtqqttttt ttqtttqcaa qcaqcaqatt acqcqcaqaa aaaaaqqatc 4860
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg 4920
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 4980
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 5040
atgettaate agtgaggeac etateteage gatetgteta tttegtteat ceatagttge 5100
ctgactcggg ggggggggg gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata 5160
ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5220
tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacqq aacqqtctqc 5280
gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca 5340
aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 5400
tetgattaga aaaacteate gageateaaa tgaaactgea atttatteat ateaggatta 5460
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 5520
ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 5580
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 5640
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 5700
ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 5760
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 5820
ategaatgea aceggegeag gaacactgee agegeateaa caatatttte acetgaatea 5880
ggatattett etaataeetg gaatgetgtt tteeegggga tegeagtggt gagtaaceat 5940
gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6000
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6060
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6120
ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt ggaatttaat 6180
egeggeeteg ageaagaegt thecegitga atatggetea taacacceet tgtattactg 6240
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6300
catcagagat tttgagacac aacgtggctt tcccccccc cccattattg aagcatttat 6360
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6420
ggggttccqc qcacatttcc ccqaaaaqtq ccacctqacq tctaaqaaac cattattatc 6480
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc
                                                                6526
```

```
<210> 138
```

<220>

<sup>&</sup>lt;211> 6607

<sup>&</sup>lt;212> DNA

<sup>&</sup>lt;213> Artificial Sequence

<223> plasmid pVR1012x/s containing HIV genes

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaqattqq 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcqqt tttggcaqta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccqcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgac acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgeggtge tgttaaeggt ggagggeagt gtagtetgag eagtaetegt tgetgeegeg 1800 egegecacca gacataatag etgacagact aacagactgt teetttecat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040 gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100 acgacegtga agttcacaac gtttgggcta cccaegcttg cgttccgacc gaccegaacc 2160 cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220 ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280 aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340 ccaacgacat gaacaaagaa atccgtaact gctccttcgc atgcatcaac tgcaacgctt 2400 ccaccatcac ccaggettgc ccgaaagtta acttcgaccc qatcccgatc cactactgcq 2460 ctccggctgg ttacgctatc ctgaaatgca acaacaaaac cttctccqqt aaaqqtccqt 2520 gcaacaacgt ttccaccgtt cagtgcaccc atggtatcaa accggttgtt tccacccage 2580 tgctgctgaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2640 acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcgtt tgcacccgtc 2700 cgaacaacaa cacccgtaaa tccatgcgta tcggtccggg tcagaccttc tacgctaccg 2760 gtgacatcat cggtgacatc cgtcaggctt actgcaacat ctccggttcc aaatggaacg 2820 aaaccctgaa acgtgttaaa gaaaaactgc aggaaaacta caacaacaac aaaaccatca 2880 aattegetee gteeteeggt ggtgacetgg aaateaceae ceacteette aactgeeqtg 2940 gtgaattett etactgeaac accgetagee tgeegtgeeg tateaacag ateateaca 3000 tgtggcaggg tgttggtcgt gctatgtacg ctccgccgat cgctggtaac atcacctgca 3060 aatccaacat caccggtctg ctgctggttc gtgacggtgg tgaagacaac aaaaccgaag 3120

aaatotteeg teegggtggt ggtaacatga aagacaactg gegtteegaa etgtacaaat 3180 acaaagttat cgaactgaaa ccgctgggta tcgctccgac cggtgctaag cttaccgttc 3240 aggetegtea getgetgtee tecategtte ageageagte caacetgetg egtgetateg 3300 aagctcagca gcacatgctg cagctgaccg tttggggtat caaacagctg cagacccgtg 3360 ttctggctat cgaacgttac ctgaaagacc agcagctcga gatctggaac aacatgacct 3420 ggatggaatg ggaccgtgaa atctccaact acaccgacac catctaccgt ctgctggaag 3480 actoccagac ccagcaggaa aaaaacgaaa aagacctgot ggototggac tootggaaaa 3540 acctgtggtc ctggttcgac atctccaact ggctgtgqtq aggatccaga tctgctqtqc 3600 cttctagttg ccagccatct gttgtttgcc cctccccgt gccttccttg accctggaag 3660 gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcat tgtctgagta 3720 ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag gattgggaag 3780 acaatagcag gcatgctggg gatgcggtgg gctctatggg tacccaggtg ctgaagaatt 3840 gacceggttc ctcctgggcc agaaagaagc aggcacatcc ccttctctgt gacacaccct 3900 gtccacgccc ctggttctta gttccagccc cactcatagg acactcatag ctcaggaggg 3960 ctccgccttc aatcccaccc gctaaagtac ttggagcggt ctctccctcc ctcatcaqcc 4020 caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata ggctattaag 4080 tgcagaggga gagaaaatgc ctccaacatg tgaggaagta atgagagaaa tcatagaatt 4140 ttaaggccat gatttaaggc catcatggcc ttaatcttcc gcttcctcgc tcactgactc 4200 getgegeteg gtegttegge tgeggegage ggtateaget caetcaaagg eggtaataeg 4260 gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag gccagcaaaa 4320 ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 4380 cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 4440 ataccaggeg tttececetg gaageteeet egtgegetet cetgtteega eeetgeeget 4500 taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 4560 ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 4620 ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 4680 aagacacgac ttategecac tggcagcagc cactggtaac aggattagca gagcgaggta 4740 tgtaggeggt getacagagt tettgaagtg gtggeetaac taeggetaca etagaagaac 4800 agtatttggt atotgogoto tgotgaagoo agttacotto ggaaaaagag ttggtagoto 4860 ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 4920 tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 4980 tragtggaar gaaaactrar gttaagggat tttggtratg agattatraa aaaggatctt 5040 cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 5100 aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5160 atttegttea tecatagttg cetgactegg ggggggggg egetgaggte tgeetegtga 5220 agaaggtgtt gctgactcat accaggcctg aatcgcccca tcatccagcc agaaagtgag 5280 ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt tgaacttttg 5340 ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatcct tcaactcagc 5400 aaaagttcga tttattcaac aaagccgccg tcccgtcaag tcagcgtaat gctctgccag 5460 tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5520 aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa 5580 ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt 5640 cegactegte caacateaat acaacetatt aattteceet egteaaaaat aaggttatea 5700 agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag cttatgcatt 5760 tetttecaga ettgtteaac aggecageca ttaegetegt cateaaaate actegeatea 5820 accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 5880 aaaggacaat tacaaacagg aatcgaatge aaccggcgca ggaacactgc cagcgcatca 5940 acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 6000 atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaatgctt gatggtcgga 6060 agaggcataa atteegteag ceagtttagt etgaceatet catetgtaae ateattggea 6120 acgetacett tgecatgitt cagaaacaac tetggegeat egggetteee atacaatega 6180 tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca 6240 geatecatgt tggaatttaa tegeggeete gageaagaeg ttteeegttg aatatggete 6300 ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6360 tttttatett gtgcaatgta acatcagaga ttttgagaca caacgtggct ttccccccc 6420

```
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6480 atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6540 gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6600 tttcgtc 6607
```

<210> 139 <211> 6676 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 139 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta ceqecatqtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 etcegeecca ttgaegeaaa tgggeggtag gegtgtaegg tgggaggtet atataaqeaq 900 agetegttta gtgaaccgte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tettatgeat getatactqt ttttqqettq qqqeetatac acceeqett cettatqeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt teettbecat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040 gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccqac accaaaqcct 2100 acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160 cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgq 2220 ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280 aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340

| anaaaaaaa t | gaagaaaga  | atoogtaact | actacttana | 6266266266 | annataget a | 2400              |
|-------------|------------|------------|------------|------------|-------------|-------------------|
|             |            |            |            |            | gaaatccgtg  |                   |
|             |            |            |            |            | ctgaaagaaa  |                   |
|             |            |            |            |            | tccaccatca  |                   |
|             |            |            |            |            | gctccggctg  |                   |
|             |            |            |            |            | tgcaacaacg  |                   |
|             |            |            |            |            | ctgctgctga  |                   |
|             |            |            |            |            | gacaacgtta  |                   |
| aaaccatcat  | cgttcacctg | aacaaatccg | ttgaaatcgt | ttgcacccgt | gctagcgctt  | 2820              |
|             |            |            |            |            | gaaaaactgc  |                   |
|             |            |            |            |            | ggtgacctgg  |                   |
| aaatcaccac  | ccactccttc | aactgccgtg | gtgaattctt | ctactgcaac | accacccgtc  | 3000              |
| tgttcaacaa  | caacgctacc | gaagacgaaa | ccatcaccct | gccgtgccgt | atcaaacaga  | 3060              |
| tcatcaacat  | gtggcagggt | gttggtcgtg | ctatgtacgc | tccgccgatc | gctggtaaca  | 3120              |
| tcacctgcaa  | atccaacatc | accggtctgc | tgctggttcg | tgacggtggt | gaagacaaca  | 3180 <sup>.</sup> |
|             |            |            |            |            | cgttccgaac  |                   |
|             |            |            |            |            | ggtgctaagc  |                   |
|             |            |            |            |            | aacctgctgc  |                   |
|             |            |            |            |            | aaacagctgc  |                   |
|             |            |            |            |            | atctggaaca  |                   |
|             |            |            |            |            | atctaccgtc  |                   |
|             |            |            |            |            | gctctggact  |                   |
|             |            |            |            |            | ggatccagat  |                   |
|             |            |            |            |            | ccttccttga  |                   |
|             |            |            |            |            | gcatcgcatt  |                   |
|             |            |            |            |            | aagggggagg  |                   |
|             |            |            |            |            | acccaggtgc  |                   |
|             |            |            |            |            | cttctctgtg  |                   |
|             |            |            |            |            | cactcatage  |                   |
|             |            |            |            |            | tetecetece  |                   |
|             |            |            |            |            | agcaagatag  |                   |
|             |            |            |            |            | tgagagaaat  |                   |
|             |            |            |            |            | cttcctcgct  |                   |
|             |            |            |            |            | actcaaaggc  |                   |
|             |            |            |            |            | gagcaaaagg  |                   |
|             |            |            |            |            | ataggeteeg  |                   |
|             |            |            |            |            | acccgacagg  |                   |
|             |            |            |            |            | ctgttccgac  |                   |
|             |            |            |            |            | cgctttctca  |                   |
|             |            |            |            |            | tgggctgtgt  |                   |
|             |            |            |            |            | gtcttgagtc  |                   |
|             |            |            |            |            | ggattagcag  |                   |
|             |            |            |            |            | acggctacac  |                   |
|             |            |            |            |            | gaaaaagagt  |                   |
|             |            |            |            |            | ttgtttgcaa  |                   |
|             |            |            |            |            | tttctacggg  |                   |
|             |            |            |            |            | gattatcaaa  |                   |
|             |            |            |            |            | tctaaagtat  |                   |
|             |            |            |            |            | ctatctcagc  |                   |
|             |            |            |            |            | gctgaggtct  |                   |
|             |            |            |            |            | catccagcca  |                   |
|             |            |            |            |            | tggtgatttt  |                   |
|             |            |            |            |            | tctgatcctt  |                   |
|             |            |            |            |            | cagcgtaatg  |                   |
|             |            |            |            |            | gagcatcaaa  |                   |
|             |            |            |            |            | aagccgtttc  |                   |
|             |            | 23         |            | <b>-</b>   | 2 2 2 2.    |                   |

```
tgtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg 5700
tetgegatte egactegtee aacateaata caacetatta attteecete gteaaaaata 5760
aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc 5820
ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca 5880
ctcgcatcaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg aaatacgcga 5940
togotyttaa aaggacaatt acaaacagga atogaatgca accggcgcag gaacactgcc 6000
agggcatcaa caatattttc acctgaatca ggatattctt ctaatacctg gaatgctgtt 6060
ttcccqqqqa tcgcagtggt gagtaaccat gcatcatcag gagtacggat aaaatgcttg 6120
atqqtcqqaa qaggcataaa ttccqtcagc cagtttagtc tgaccatctc atctqtaaca 6180
tcattggcaa cgctaccttt gccatgtttc agaaacaact ctggcgcatc gggcttccca 6240
tacaatcgat agattgtcgc acctgattgc ccgacattat cgcgagccca tttataccca 6300
tataaatcag catccatgtt ggaatttaat cgcggcctcg agcaagacgt ttcccgttga 6360
atatqqctca taacacccct tgtattactg tttatgtaag cagacagttt tattgttcat 6420
qatqatatat ttttatcttq tqcaatqtaa catcagagat tttgagacac aacgtggctt 6480
tcccccccc cccattattg aagcatttat cagggttatt gtctcatgag cggatacata 6540
tttqaatqta tttaqaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 6600
ccacctqacq tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc 6660
acgaggccct ttcgtc
                                                                  6676
```

<210> 140 <211> 6637 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 140 tegegegttt eggtgatgae ggtgaaaade tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cocqcotqqc tqaccqccca acqacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgce aagagtgacg taagtacege etatagacte tataggcaca eccetttggc 1080 tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgac aeggactetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500

ageggeteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 qcaqaaqaaq atqcaqqcaq ctqaqttqtt qtattctqat aaqaqtcaqa qqtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgtqt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040 qtqttccqqt ttggaccqac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100 acquecqtqu agttcacauc gtttgggctu cccacgcttg cgttccgacc gacccgaacc 2160 cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220 ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280 aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340 ccaacgacat gaacaaagaa atccgtaact gctccttcaa caccacca gaaatccgtg 2400 acaaaaaaca gcagggttac gctctgttct accgtccgga catcgttctg ctgaaagaaa 2460 accetaacaa ctccaacaac tccgaataca tcctgatcaa ctgcaacgct tccaccatca 2520 cccaggettg cccgaaagtt aacttcgacc cgatcccgat ccactactgc geteeggetg 2580 gttacgctat cctgaaatgc aacaacaaaa ccttctccgg taaaggtccg tgcaacaacg 2640 tttccaccqt tcaqtqcacc catqqtatca aaccqqttqt ttccacccaq ctqctqca 2700 acggttccct ggctgaaaaa gaaatcatca tccgttccga aaacctgacc gacaacgtta 2760 aaaccatcat cgttcacctg aacaaatccg ttgaaatcgt ttgcacccgt gctagcgctt 2820 actgcaacat ctccggttcc aaatggaacg aaaccctgaa acgtgttaaa gaaaaactgc 2880 aggaaaacta caacaacaac aaaaccatca aattcgctcc gtcctccggt ggtgacctgg 2940 aaatcaccac ccactccttc aactgccgtg gtgaattctt ctactgcaac accgctagcc 3000 tgccgtgccg tatcaaacag atcatcaaca tgtggcaggg tgttggtcgt gctatgtacg 3060 ctccgccgat cgctggtaac atcacctgca aatccaacat caccggtctg ctgctggttc 3120 qtqacqqtqq tqaaqacaac aaaaccgaag aaatcttccg tccgggtggt ggtaacatga 3180 aaqacaactq qcqttccqaa ctqtacaaat acaaagttat cgaactgaaa ccgctgggta 3240 tegeteegae eggtgetaag ettacegtte aggetegtea getgetgtee tecategtte 3300° aqcaqcaqte caacetgetg egtgetateg aageteagea geacatgetg cagetgaceg 3360 tttqqqqtat caaacagctg cagacccgtg ttctggctat cgaacgttac ctgaaagacc 3420 agcagetega gatetggaae aacatgaeet ggatggaatg ggaeegtgaa atetecaaet 3480 acaccgacac catctaccgt ctgctggaag actcccagac ccagcaggaa aaaaacgaaa 3540 aagacetget ggetetggae teetggaaaa acetgtggte etggttegae atetecaaet 3600 ggctgtggtg aggatccaga tctgctgtgc cttctagttg ccagccatct gttgtttgcc 3660 cctcccccqt gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa 3720 qqcaqqacaq caaqqqqqaq qattqqqaaq acaataqcaq qcatqctqqq gatqcqqtqq 3840 qctctatqqq tacccaqqtq ctqaaqaatt qacccqqttc ctcctqqqcc agaaagaagc 3900 aggcacatec cettetetgt gacacaceet gtecaegeee etggttetta gttecageee 3960 cactcatagg acactcatag ctcaggaggg ctccgccttc aatcccaccc gctaaagtac 4020 ttggagcggt ctctccctcc ctcatcagcc caccaaacca aacctagcct ccaagagtgg 4080 gaagaaatta aagcaagata ggctattaag tgcagaggga gagaaaatgc ctccaacatg 4140 tgaggaagta atgagagaaa tcatagaatt ttaaggccat gatttaaggc catcatggcc 4200 ttaatcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc 4260 ggtatcaget cacteaaagg eggtaataeg gttatceaea gaateagggg ataaegeagg 4320 aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 4380 ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4440 qaqqtqqcqa aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4500 egtgegetet cetgtteega ecetgeeget taceggatac etgteegeet ttetecette 4560 gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt 4620 tegetecaag etgggetgtg tgeaegaace eeegtteag eeegaeeget gegeettate 4680 eggtaactat egtettgagt ecaaceeggt aagacaegae ttategeeae tggeageage 4740 cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 4800

```
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 4860
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4920
cqqtqqtttt tttqtttqca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 4980
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 5040
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5100
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 5160
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5220
ggggggggg cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat accaggcctg 5280
aatcqccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttgttgta 5340
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcggg 5400
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5460
tecegteaag teagegtaat getetgeeag tgttacaace aattaaceaa ttetgattag 5520
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5580
tatttttqaa aaaqccqttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5640
atggcaagat cotggtatcg gtotgcgatt cogactogto caacatcaat acaacctatt 5700
aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5760
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5820
ttacqctcqt catcaaaatc actcqcatca accaaaccqt tattcattcq tgattqcqcc 5880
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 5940
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 6000
tctaatacct ggaatgctqt tttcccqqqq atcqcaqtqq tgagtaacca tgcatcatca 6060
qqaqtacqqa taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6120
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6180
tetggegeat egggetteec atacaatega tagattgteg cacetgattg ecegacatta 6240
tcgcgagccc atttataccc atataaatca gcatccatgt tggaatttaa tcgcggcctc 6300
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6360
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6420
ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta tcagggttat 6480
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6540
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6600
                                                                  6637
acctataaaa ataggcgtat cacgaggccc tttcgtc
```

```
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 141
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggeege teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta cccccatqtt qacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
```

<210> 141 <211> 6718

| catcaatggg | cgtggatagc | ggtttgactc | acggggattt | ccaagtctcc | accccattga | 780  |
|------------|------------|------------|------------|------------|------------|------|
|            | agtttgtttt |            |            |            |            |      |
|            | ttgacgcaaa |            |            |            |            |      |
|            | gtgaaccgtc |            |            |            |            |      |
|            | cgggaccgat |            |            |            |            |      |
|            | aagagtgacg |            |            |            |            |      |
|            | gctatactgt |            |            |            |            |      |
|            | tatagettag |            |            |            |            |      |
|            | gatactttcc |            |            |            |            |      |
|            | atgccaatac |            |            |            |            |      |
|            | ccatttatta |            |            |            |            |      |
|            | attaaacata |            |            |            |            |      |
|            | tctccggtag |            |            |            |            |      |
|            | ggtcgctcgg |            |            |            |            |      |
|            | ccaccaccac |            |            |            |            |      |
|            | gtggagattg |            |            |            |            |      |
|            | atgcaggcag |            |            |            |            |      |
|            | tgttaacggt |            |            |            |            |      |
|            | gacataatag |            |            |            |            |      |
|            | gtcgtcgaca |            |            |            |            |      |
|            | tcgtggtatc |            |            |            |            |      |
|            | gatcatcatc |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | ttggaccgac |            |            |            |            |      |
|            | agttcacaac |            |            |            |            |      |
|            | cgttctggaa |            |            |            |            |      |
|            | gcacgaagac |            |            |            |            |      |
|            | getgtgegtt |            |            |            |            |      |
|            | gaacaaagaa |            |            |            |            |      |
|            | gcagggttac |            |            |            |            |      |
|            | ctccaacaac |            |            |            |            |      |
|            | cccgaaagtt |            |            |            |            |      |
|            | cctgaaatgc |            |            |            |            |      |
|            | tcagtgcacc |            |            |            |            |      |
|            | ggctgaaaaa |            |            |            |            |      |
|            | cgttcacctg |            |            |            |            |      |
|            | atccatgcgt |            |            |            |            |      |
|            | ccgtcaggct |            |            |            |            |      |
|            | agaaaaactg |            |            |            |            |      |
|            | tggtgacctg |            |            |            |            |      |
|            | caccgctagc |            |            |            |            |      |
|            | tgctatgtac |            |            |            |            |      |
|            | gctgctggtt |            |            |            |            |      |
|            | tggtaacatg |            |            |            |            |      |
|            | accgctgggt |            |            |            |            |      |
|            | ctccatcgtt |            |            |            |            |      |
|            | gcagctgacc |            |            |            |            |      |
|            | cctgaaagac |            |            |            |            |      |
|            | aatctccaac |            |            |            |            |      |
|            | aaaaaacgaa |            |            |            |            |      |
|            | catctccaac |            |            |            |            |      |
|            | tgttgtttgc |            |            |            |            |      |
|            | ttcctaataa |            |            |            |            |      |
|            | aaaraaaara |            |            |            |            |      |
|            | ggatgcggtg |            |            |            |            |      |
| cctcctgggc | cagaaagaag | caggcacatc | cccttctctg | tgacacaccc | tgtccacgcc | 4020 |

```
cetggttett agttecagee ceacteatag gacacteata geteaggagg geteegeett 4080
caatcccacc cgctaaagta cttggagcgg tctctccctc cctcatcaqc ccaccaaacc 4140
aaacctagcc tccaagagtg ggaagaaatt aaagcaagat aggctattaa gtgcagaggg 4200
agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat tttaaggcca 4260
tgatttaagg ccatcatggc cttaatcttc cgcttcctcg ctcactgact cgctgcgctc 4320
ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac 4380
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa qgccaqcaaa aqqccaqqaa 4440
cogtaaaaag googogttgo tggogttttt coataggoto cgccccctg acgagcatca 4500
caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa qataccaqqc 4560
gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accetgccgc ttaccggata 4620
cetgtcegcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 4680
teteagtteg gtgtaggteg ttegeteeaa getgggetgt gtgcacqaac cecceqttea 4740
gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga 4800
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4860
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4920
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4980
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcaqcaqa ttacqcqcaq 5040
aaaaaaagga totcaagaag atootttgat ottttotacg gggtotgacg otcagtggaa 5100
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 5160
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5220
tgacagttac caatgettaa teagtgagge acetatetea gegatetgte tatttegtte 5280
atccatagtt gcctgactcg ggggggggg gcgctgaggt ctgcctcgtg aagaaqqtqt 5340
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5400
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt qctttqccac 5460
ggaacggtot gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag caaaagttcg 5520
atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5580
caattaacca attetgatta gaaaaactca tegageatca aatgaaactg caatttatte 5640
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5700
tcaccgagge agttccatag gatggcaaga tcctggtatc ggtctgcgat tccgactcgt 5760
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5820
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5880
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 5940
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaqqacaa 6000
ttacaaacag gaatcgaatg caaccggcgc aggaacactq ccaqcqcatc aacaatattt 6060
teacetgaat caggatatte ttetaatace tggaatgetg tttteeeggg gategeagtg 6120
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6180
aattoogtoa gooagtttag totgacoato toatotgtaa catcattggo aacgotacot 6240
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6300
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6360
ttggaabtta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6420
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6480
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttcccccc cccccattat 6540
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6600
aataaacaaa taggggttoc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6660
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc
```

```
<210> 142
<211> 6790
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HTV genes
```

<400> 142 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccaqtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 etecgececa ttgacgeaaa tgggeggtag gegtgtaegg tgggaggtet atataageag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getataetgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecce 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tectttecat gggtetttte 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040 gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100 acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160 cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220 ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280 aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagccgtaac tgctccttca 2340 acaccaccac cgaaatccgt gacaaaaaac agcagggtta cgctctgttc taccgtccgg 2400 acategttet getgaaagaa aacegtaaca actecaacaa eteegaatae ateetgatea 2460 actgcaacge ttecaccate acceaggett geoegaaagt taacttegae eegateeega 2520 tecactactg egeteegget ggttacgeta teetgaaatg caacaacaaa acetteteeg 2580 gtaaaggtcc gtgcaacaac gtttccaccg ttcagtgcac ccatggtatc aaaccggttg 2640 tttccaccca gctgctgctg aacggttccc tggctgaaaa agaaatcatc atccgttccg 2700 aaaacctgac cqacaacgtt aaaaccatca tcgttcacct qaacaaatcc qttqaaatcg 2760 tttgcacccg tccgaacaac aacacccgta aatccatgcg tatcggtccg ggtcagacct 2820 totacgotac oggtgacato atoggtgaca tocgtcaggo ttactgoaac atotocggtt 2880 ccaaatggaa cgaaaccctg aaacgtgtta aagaaaaact gcaggaaaac tacaacaaca 2940 acaaaaccat caaatteget cegteeteeg gtggtgacet ggaaateace acceaeteet 3000 tcaactgccg tggtgaattc ttctactgca acaccacccg tctgttcaac aacaacgcta 3060 ccgaagacga aaccatcacc ctgccgtgcc gtatcaaaca gatcatcaac atgtggcagg 3120 gtgttggtcg tgctatgtac gctccgccga tcgctggtaa catcacctgc aaatccaaca 3180

| tcaccggtct | gctgctggtt | cgtgacqqtq | gtgaagacaa  | caaaaccgaa | gaaatettee        | 3240 |
|------------|------------|------------|-------------|------------|-------------------|------|
| gtccgggtgg | tggtaacatg | aaaqacaact | ggcgttccga  | actgtacaaa | tacaaagtta        | 3300 |
| tcgaactgaa | accgctgggt | atcqctccqa | ccaatactaa  | gcttaccgtt | cagactcatc        | 3360 |
| agctgctgtc | ctccatcqtt | cagcagcagt | ccaacctoct  | gcgtgctatc | gaageteage        | 3420 |
| agcacatgct | gcagetgace | atttagaata | tcaaacagct  | gcagacccgt | gttctggcta        | 3480 |
| tcqaacqtta | cctgaaagac | cagcageteg | agatctggaa  | caacatgacc | tagatagaat        | 3540 |
| gggaccgtga | aatctccaac | tacaccgaca | ccatctacca  | tctgctggaa | gactcccaga        | 3600 |
| cccaqcaqqa | aaaaaacgaa | aaagacctgc | taactotaaa  | ctcctggaaa | aacctataat        | 3660 |
| cctggttcga | catctccaac | taactataat | acatcaaaat  | cttcatcatg | atcottooto        | 3720 |
| gtctgatcgg | tctgcgtatc | atcttcocto | ttctatccat  | ctgaggatcc | agatetgetg        | 3780 |
| taccttctag | ttaccaacca | tctattatt  | acceptede   | cgtgccttcc | traaccetag        | 3840 |
| aaggtgccac | teccaetate | ctttcctaat | aaaatgagga  | aattgcatcg | cattetetea        | 3900 |
| gtaggtgtca | ttctattctg | gagaataaaa | tagagagaga  | cagcaagggg | gaggattgg         | 3960 |
| aagacaatag | caggcatgct | aggastacaa | tagactetat  | gggtacccag | gaggactggg        | 4020 |
| attgacccgg | ttcctcctaa | accadaaada | agraggrada  | tccccttctc | tatasassa         | 4020 |
| cctatccaca | cccctaattc | ttaattccaa | cccactcat   | aggacactca | tageterage        | 4140 |
| agactccacg | ttcaatccca | cccachaaaa | tacttogage  | ggtctctccc | tageteatga        | 4140 |
| gcccaccaaa | ccaaaccted | cctccaaga  | tacceggage  | thanagana  | ataggetate        | 4200 |
| geccaceaaa | ggagagaaaa | tacatacasa | cygyaayaaa  | ttaaagcaag | acaggetatt        | 4260 |
| attttaagg  | ggagagaaaa | ggggatgatg | acgigaggaa  | gtaatgagag | aaatcataga        | 4320 |
| accccaagge | tacgacctaa | ggccaccacg | geettaatet  | tecgetteet | cgcccactga        | 4380 |
| aggettatag | acagaataa  | ggetgeggeg | ageggeatea  | gctcactcaa | aggcggtaat        | 4440 |
| acggccaccc | acagaatcag | gggataacgc | aggaaagaac  | atgtgagcaa | aaggccagca        | 4500 |
| aaaggccagg | aaccgcaaaa | aggeegeget | getggegttt  | ttccataggc | teegeeeeee        | 4560 |
| cgacgagcat | cacaaaaacc | gacgcccaag | tcagaggtgg  | cgaaacccga | caggactata        | 4620 |
| aagataccag | gcgtttcccc | ctggaagete | cctcgtgcgc  | tctcctgttc | cgaccctgcc        | 4680 |
| gcttaccgga | tacctgtccg | catttetece | ttcgggaagc  | gtggcgcttt | ctcatagctc        | 4740 |
| acgetgtagg | tateteaget | cggtgtaggt | cgttcgctcc  | aagctgggct | gtgtgcacga        | 4800 |
| acccccgtt  | cagecegaee | getgegeett | atccggtaac  | tatcgtcttg | agtccaaccc        | 4860 |
| ggtaagacac | gacttatege | cactggcagc | agccactggt  | aacaggatta | gcagagcgag        | 4920 |
| gtatgtaggc | ggtgctacag | agttcttgaa | gtggtggcct  | aactacggct | acactagaag        | 4980 |
| aacagtattt | ggtatctgcg | ctctgctgaa | gccagttacc  | ttcggaaaaa | gagttggtag        | 5040 |
| ctcttgatcc | ggcaaacaaa | ccaccgctgg | tagcggtggt  | ttttttgttt | gcaagcagca        | 5100 |
| gattacgcgc | agaaaaaaag | gatctcaaga | agateetttg  | atcttttcta | cggggtctga        | 5160 |
| cgctcagtgg | aacgaaaact | cacgttaagg | gattttggtc  | atgagattat | caaaaaggat        | 5220 |
| cttcacctag | atcettttaa | attaaaaatg | aagttttaaa  | tcaatctaaa | gtatatatga        | 5280 |
| gtaaacttgg | tctgacagtt | accaatgett | aatcagtgag  | gcacctatct | cagcgatctg        | 5340 |
| tctatttcgt | tcatccatag | ttgcctgact | caaaaaaaaaa | gggcgctgag | gtctgcctcg        | 5400 |
| tgaagaaggt | gttgctgact | cataccagge | ctgaatcgcc  | ccatcatcca | gccagaaagt        | 5460 |
| gagggagcca | cggttgatga | gagctttgtt | gtaggtggac  | cagttggtga | <u>ttttgaactt</u> | 5520 |
| ttgctttgcc | acggaacggt | ctgcgttgtc | gggaagatgc  | gtgatctgat | ccttcaactc        | 5580 |
| agcaaaagtt | cgatttattc | aacaaagccg | ccgtcccgtc  | aagtcagcgt | aatgctctgc        | 5640 |
| cagtgttaca | accaattaac | caattctgat | tagaaaaact  | catcgagcat | caaatgaaac        | 5700 |
| tgcaatttat | tcatatcagg | attatcaata | ccatattttt  | gaaaaagccg | tttctgtaat        | 5760 |
| gaaggagaaa | actcaccgag | gcagttccat | aggatggcaa  | gatcctggta | tcggtctgcg        | 5820 |
| attccgactc | gtccaacatc | aatacaacct | attaatttcc  | cctcgtcaaa | aataaggtta        | 5880 |
| tcaagtgaga | aatcaccatg | agtgacgact | gaatccggtg  | agaatggcaa | aagcttatgc        | 5940 |
| atttctttcc | agacttgttc | aacaggccag | ccattacgct  | cgtcatcaaa | atcactcgca        | 6000 |
| tcaaccaaac | cgttattcat | tcgtgattgc | gcctgagcga  | gacgaaatac | gcgatcgctg        | 6060 |
| ttaaaaggac | aattacaaac | aggaatcgaa | tgcaaccggc  | gcaggaacac | tgccagcgca        | 6120 |
| tcaacaatat | tttcacctga | atcaggatat | tcttctaata  | cctggaatgc | tgttttcccg        | 6180 |
| gggatcgcag | tggtgagtaa | ccatgcatca | tcaggagtac  | ggataaaatg | cttgatggtc        | 6240 |
| ggaagaggca | taaattccgt | cagccagttt | agtctgacca  | tctcatctgt | aacatcattg        | 6300 |
| gcaacgctac | ctttgccatg | tttcagaaac | aactctggcg  | catcgggctt | cccatacaat        | 6360 |
| cgatagattg | tcgcacctga | ttgcccgaca | ttatcgcgag  | cccatttata | cccatataaa        | 6420 |
| tcagcatcca | tgttggaatt | taatcgcggc | ctcgagcaag  | acgtttcccg | ttgaatatgg        | 6480 |
|            |            |            |             | _          |                   |      |

```
ctcataacac cccttgtatt actgtttatg taagcagaca gttttattgt tcatgatgat 6540 atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttcccc 6600 ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6660 tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct 6720 gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6780 ccctttcgtc 6720
```

<210> 143
<211> 6658
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 143

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeq qaqaeqqtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaqattgq 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cocgcotggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatqttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tottatgcat gotatactgt ttttggottg gggootatac accoccgott cottatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct totocggtag cggcggagct tocacatccg agccctggtc ccatgcctcc 1500 agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcqt tgctgccgcq 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ceatgegtgt tegtggtate eegegtaact ggeegeagtg gtggatgtgg ggtateetgg 1980 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040 gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100 acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160 cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220

|            |            |            |            | , ,        |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | gcacgaagac |            |            |            |            |      |
|            | gctgtgcgtt |            |            |            |            |      |
| ccaccatcac | ccaggcttgc | ccgaaagtta | acttcgaccc | gatcccgatc | cactactgcg | 2400 |
|            | ttacgctatc |            |            |            |            |      |
| gcaacaacgt | ttccaccgtt | cagtgcaccc | atggtatcaa | accggttgtt | tccacccagc | 2520 |
| tgctgctgaa | cggttccctg | gctgaaaaag | aaatcatcat | ccgttccgaa | aacctgaccg | 2580 |
|            | aaccatcatc |            |            |            |            |      |
|            | cacccgtaaa |            |            |            |            |      |
|            | cggtgacatc |            |            |            |            |      |
|            | acgtgttaaa |            |            |            |            |      |
|            | gtcctccggt |            |            |            |            |      |
|            | ctactgcaac |            |            |            |            |      |
|            | gccgtgccgt |            |            |            |            |      |
|            | tccgccgatc |            |            |            | -          |      |
|            | tgacggtggt |            |            |            |            |      |
|            | agacaactgg |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | cgctccgacc |            |            |            |            |      |
|            | gcagcagtcc |            |            |            |            |      |
|            | ttggggtatc |            |            |            |            |      |
|            | gcagctcgag |            |            |            |            |      |
|            | caccgacacc |            |            |            |            |      |
|            | agacctgctg |            |            |            |            |      |
|            | gctgtggtac |            |            |            |            |      |
|            | cttcgctgtt |            |            |            |            |      |
|            | tgttgtttgc |            |            |            |            |      |
| ccactgtcct | ttcctaataa | aatgaggaaa | ttgcatcgca | ttgtctgagt | aggtgtcatt | 3780 |
| ctattctggg | gggtggggtg | gggcaggaca | gcaaggggga | ggattgggaa | gacaatagca | 3840 |
| ggcatgctgg | ggatgcggtg | ggctctatgg | gtacccaggt | gctgaagaat | tgacccggtt | 3900 |
| cctcctgggc | cagaaagaag | caggcacatc | cccttctctg | tgacacaccc | tgtccacgcc | 3960 |
| cctggttctt | agttccagcc | ccactcatag | gacactcata | gctcaggagg | gctccgcctt | 4020 |
|            | cgctaaagta |            |            |            |            |      |
|            | tccaagagtg |            |            |            |            |      |
|            | cctccaacat |            |            |            |            |      |
|            | ccatcatggc |            |            |            |            |      |
|            | ctgcggcgag |            |            |            |            |      |
|            | gataacgcag |            |            |            |            |      |
|            | gccgcgttgc |            |            |            |            |      |
|            | cgctcaagtc |            |            |            |            |      |
|            | ggaagctccc |            |            |            |            |      |
|            | tttctccctt |            |            |            |            |      |
|            | gtgtaggtcg |            |            |            |            |      |
|            | tgcgccttat |            |            |            |            |      |
|            | ctggcagcag |            |            |            |            |      |
|            | ttcttgaagt |            |            |            |            |      |
|            | ctgctgaagc |            |            |            |            |      |
|            | accgctggta |            |            |            |            |      |
|            |            |            |            |            |            |      |
|            | tctcaagaag |            |            |            |            |      |
|            | cgttaaggga |            |            |            | -          |      |
|            | taaaaatgaa |            |            |            |            |      |
|            | caatgcttaa |            |            |            |            |      |
|            | gcctgactcg |            |            |            |            |      |
|            | taccaggeet |            |            |            |            |      |
|            | gctttgttgt |            |            |            |            |      |
|            | gcgttgtcgg |            |            |            |            |      |
| atttattcaa | caaagccgcc | gtcccgtcaa | gtcagcgtaa | tgctctgcca | gtgttacaac | 5520 |
|            |            |            |            |            |            |      |

```
caattaacca attotgatta gaaaaactca togagcatca aatgaaactg caatttatto 5580
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5640
tcaccgaggc agttccatag gatggcaaga tcctggtatc ggtctgcgat tccgactcgt 5700
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5760
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5820
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 5880
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 5940
ttacaaacaq gaatcqaatq caaccggcgc aggaacactg ccaqcgcatc aacaatattt 6000
tcacctqaat caggatattc ttctaatacc tggaatgctg ttttcccggg gatcgcagtg 6060
qtqaqtaacc atqcatcatc aqqaqtacqq ataaaatqct tqatqqtcqq aaqaqqcata 6120
aattoogtoa gooagtttag totgacoato toatotgtaa catcattggo aacgotacot 6180
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6240
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6300
ttqqaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6360
cttqtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6420
tqtqcaatqt aacatcagaq attttgagac acaacqtgqc tttccccccc cccccattat 6480
tqaaqcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6540
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctqa cgtctaagaa 6600
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc
```

<210> 144 <211> 6577 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 144 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattae 360 qqqqtcatta qttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccqcctqqc tqaccqccca acqacccccq cccattqacq tcaataatqa cqtatqttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tqcccacttq qcaqtacatc aagtgtatca tatgccaagt acqcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eeeetttgge 1080 tettatgeat getatactgt tittggettg gggeetatae acceeegett cettatgeta 1140

taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440

| catgggctct | tctccggtag | cggcggagct | tccacatccq  | agccctggtc | ccatgcctcc | 1500 |
|------------|------------|------------|-------------|------------|------------|------|
| agcggctcat | ggtcgctcgg | cagctccttq | ctcctaacag  | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg  | tggcggtagg | gratgratch | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg | actgacgcag  | atggaagact | taaggragge | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat  | aagagtcaga | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag  | cagtactcgt | tactaccaca | 1900 |
| cqcqccacca | gacataatag | ctgacagact | aacagactgt  | tcctttccat | agatettte  | 1060 |
| tgcagtcacc | gtcgtcgaca | catataatca | gatatogogo  | ccgctctaga | gageeeeee  | 1000 |
| ccatacatat | tcataatata | ccacataact | aaccacaata  | gtggatgtgg | gatategeea | 1000 |
| gtttctggat | gatcatcatc | taccatatta | ttootaacat  | gtgggttacc | ggtaccccgg | 1200 |
| gtattccaat | ttggaccgac | gctaaaacca | ccctattcta  | cgcttccgac | accasages  | 2040 |
| acgaccgtga | agttcacaac | attragacta | cccacactta  | cgttccgacc | accadageee | 2100 |
| cacaggaaat | cattetagaa | aacottacco | aaaacttcaa  | catgtggaaa | gacccgaacc | 5100 |
| ttgaccagat | acacaaaaa  | atcatctccc | tataaacccaa | gtccctgaaa | aacyacatgg | 2220 |
| aactgacccc | actatacatt | accetocact | caccaacca   | tagcatcaac | tagasasata | 2280 |
| ccaccatcac | ccaggettec | ccassatts  | acttogacco  | gatecegate | cgcaacgctt | 2340 |
| ctccaactaa | ttacgctate | ctgaaatgca | accecgaece  | gateeegate | cactactgcg | 2400 |
| gcaacaacgt | ttccaccatt | cagaaacgca | acaacaaaac  | cttctccggt | aaaggteegt | 2460 |
| tactactase | coatteesta | ggtgcaccc  | acyguateaa  | accggttgtt | tccacccage | 2520 |
| acaacattaa | aggetteetg | gergaaaaag | aaaccaccac  | ccgttccgaa | aacctgaccg | 2580 |
| staggetta  | atcoaccate | gereaceega | acaaatccgt  | tgaaatcgtt | tgcacccgtg | 2640 |
| ccagegeeta | crycaacacc | teeggtteea | aacggaacga  | aaccctgaaa | cgtgttaaag | 2700 |
| adadactyca | ggaaaactac | aacaacaaca | aaaccaccaa  | attcgctccg | tcctccggtg | 2760 |
| grgaeergga | aatcaccacc | cactcettca | actgccgtgg  | tgaattette | tactgcaaca | 2820 |
| ccaccegtet | gttcaacaac | aacgctaccg | aagacgaaac  | catcaccctg | ccgtgccgta | 2880 |
| tcaaacagat | catcaacatg | tggcagggtg | ttggtcgtgc  | tatgtacgct | ccgccgatcg | 2940 |
| ctggtaacat | cacctgcaaa | tccaacatca | ccggtctgct  | gctggttcgt | gacggtggtg | 3000 |
| aagacaacaa | aaccgaagaa | atcttccgtc | cgggtggtgg  | taacatgaaa | gacaactggc | 3060 |
| gttccgaact | gtacaaatac | aaagttatcg | aactgaaacc  | gctgggtatc | gctccgaccg | 3120 |
| grgcraager | taccgttcag | gctcgtcagc | tgctgtcctc  | catcgttcag | cagcagtcca | 3180 |
| acctgctgcg | tgctatcgaa | gctcagcagc | acatgctgca  | gctgaccgtt | tggggtatca | 3240 |
| aacagctgca | gacccgtgtt | ctggctatcg | aacgttacct  | gaaagaccag | cagctcgaga | 3300 |
| tctggaacaa | catgacctgg | atggaatggg | accgtgaaat  | ctccaactac | accgacacca | 3360 |
| tctaccgtct | gctggaagac | tcccagaccc | agcaggaaaa  | aaacgaaaaa | gacctgctgg | 3420 |
| ctctggactc | ctggaaaaac | ctgtggtcct | ggttcgacat  | ctccaactgg | ctgtggtaca | 3480 |
| tcaaaatctt | catcatgatc | gttggtggtc | tgatcggtct  | gcgtatcatc | ttcgctgttc | 3540 |
| tgtccatctg | aggatccaga | tctgctgtgc | cttctagttg  | ccagccatct | gttgtttgcc | 3600 |
| cctccccgt  | gccttccttg | accctggaag | gtgccactcc  | cactgtcctt | tcctaataaa | 3660 |
| atgaggaaat | tgcatcgcat | tgtctgagta | ggtgtcattc  | tattctgggg | ggtggggtgg | 3720 |
| ggcaggacag | caagggggag | gattgggaag | acaatagcag  | gcatgctggg | gatgcggtgg | 3780 |
| gctctatggg | tacccaggtg | ctgaagaatt | gacccggttc  | ctcctgggcc | agaaagaagc | 3840 |
| aggcacatcc | ccttctctgt | gacacaccct | gtccacgccc  | ctggttctta | gttccagccc | 3900 |
| cactcatagg | acactcatag | ctcaggaggg | ctccgccttc  | aatcccaccc | gctaaagtac | 3960 |
| ttggagcggt | ctctccctcc | ctcatcagcc | caccaaacca  | aacctagcct | ccaagagtgg | 4020 |
| gaagaaatta | aagcaagata | ggctattaag | tgcagaggga  | gagaaaatgc | ctccaacatg | 4080 |
| tgaggaagta | atgagagaaa | tcatagaatt | ttaaggccat  | gatttaaggc | catcatggcc | 4140 |
| ttaatcttcc | gcttcctcgc | tcactgactc | gctgcgctcg  | gtcgttcggc | tgcggcgagc | 4200 |
| ggtatcagct | cactcaaagg | cggtaatacg | gttatccaca  | gaatcagggg | ataacgcagg | 4260 |
| aaagaacatg | tgagcaaaag | gccagcaaaa | ggccaggaac  | cgtaaaaagg | ccgcgttgct | 4320 |
| ggcgtttttc | cataggctcc | gcccccctga | cgagcatcac  | aaaaatcgac | gctcaagtca | 4380 |
| gaggtggcga | aacccgacag | gactataaag | ataccaggcg  | tttccccctg | gaageteect | 4440 |
| cgtgcgctct | cctgttccga | ccctgccgct | taccggatac  | ctgtccgcct | ttetecette | 4500 |
| gggaagcgtg | gcgctttctc | atagctcacg | ctgtaggtat  | ctcagttcgg | tgtaggtcgt | 4560 |
| tcgctccaag | ctgggctgtq | tgcacgaacc | ccccgttcag  | cccgaccgct | gcgccttate | 4620 |
| cggtaactat | cgtcttgagt | ccaacccggt | aagacacqac  | ttatcgccac | togcagcagc | 4680 |
| cactggtaac | aggattagca | gagcgaggta | tgtaggcggt  | gctacagagt | tcttgaagtg | 4740 |
|            | =          | .= -       |             |            |            |      |

```
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 4800
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4860
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 4920
teetttgate ttttetaegg ggtetgaege teagtggaae gaaaaeteae gttaagggat 4980
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5040
ttttaaatca atctaaagta tatatgagta aacttggtot qacagttacc aatgcttaat 5100
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcqq 5160
ggggggggg cgctgaggtc tgcctcgtga aqaaqqtqtt qctqactcat accaqqcctq 5220
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttgttqta 5280
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcggg 5340
aagatgegtg atetgateet teaacteage aaaagttega tttatteaac aaageegeeg 5400
tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa ttctgattag 5460
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaqqatt atcaatacca 5520
tatttttgaa aaagcegttt etgtaatgaa ggagaaaact cacegaggca gttecatagg 5580
atggcaagat cotggtatog gtotgcgatt cogactogto caacatcaat acaacctatt 5640
aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5700
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5760
ttacgetegt cateaaaate actegeatea aceaaacegt tatteatteg tgattgegee 5820
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacaqq aatcqaatqc 5880
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 5940
tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tgcatcatca 6000
ggagtacgga taaaatgctt gatggtcgga aqaggcataa attccqtcag ccagtttagt 6060
etgaceatet catetgtaac atcattggea acgetacett tgecatgttt caqaaacaac 6120
tetggegeat egggetteee atacaatega tagattgteg cacetgattg eeegacatta 6180
tegegageee atttatacee atataaatea geateeatgt tggaatttaa tegeggeete 6240
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6300
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6360
ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta tcagggttat 6420
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6480
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6540
acctataaaa ataggcgtat cacgaggccc tttcgtc
<210> 145
<211> 6538
```

```
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 145
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggegggtg teggggetgg ettaactatg eggcateaga geagattgta etgagagtge 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
eccgcctggc tgaccgcca acgacccccg cccattgacq tcaataatga cqtatqttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacqtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
```

<212> DNA

WO 02/32943 PCT/US01/25721 Catcaatggg cgtggatagc ggtttgactc accgggattt ccaagtctcc accccattca 780

| catcaatooo | catagataga | agtttgactc | acqqqqattt | ccaagtetee | accccattga | 780  |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | gtcgtaacaa |      |
|            |            |            |            |            | atataagcag |      |
|            |            |            |            |            | ttgacctcca |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | gaacgcggat |      |
|            |            |            |            |            | cccctttggc |      |
|            |            |            |            |            | ccttatgcta |      |
|            |            |            |            |            | accactcccc |      |
|            |            |            |            |            | caactatctc |      |
|            |            |            |            |            | tatttttaca |      |
|            |            |            |            |            | cccccgtgcc |      |
|            |            |            |            |            | gtgttccgga |      |
|            |            |            |            |            | ccatgcctcc |      |
|            |            |            |            |            | acttaggcac |      |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact | taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg | 1800 |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | tcctttccat | gggtcttttc | 1860 |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatcgcca | 1920 |
| ccatgcgtgt | tcgtggtatc | ccgcgtaact | ggccgcagtg | gtggatgtgg | ggtatcctgg | 1980 |
| gtttctggat | gatcatcatc | tgccgtgttg | ttggtaacat | gtgggttacc | gtttactacg | 2040 |
| gtgttccggt | ttggaccgac | gctaaaacca | ccctgttctg | cgcttccgac | accaaagcct | 2100 |
| acgaccgtga | agttcacaac | gtttgggcta | cccacgcttg | cgttccgacc | gacccgaacc | 2160 |
|            |            |            |            |            | aacgacatgg |      |
|            |            |            |            |            | ccgtgcgtta |      |
|            |            |            |            |            | tgcaacgctt |      |
|            |            |            |            |            | cactactgcg |      |
|            |            |            |            |            | aaaggtccgt |      |
|            |            |            |            |            | tccacccagc |      |
|            |            |            |            |            | aacctgaccg |      |
| acaacgttaa | aaccatcatc | gttcacctga | acaaatccgt | tgaaatcgtt | tgcacccgtg | 2640 |
|            |            |            |            |            | cgtgttaaag |      |
|            |            |            |            |            | tcctccggtg |      |
|            |            |            |            |            | tactgcaaca |      |
|            |            |            |            |            | gttggtcgtg |      |
|            |            |            |            |            | accggtctgc |      |
|            |            |            |            |            | ccgggtggtg |      |
|            |            |            |            |            | gaactgaaac |      |
|            |            |            |            |            | ctgctgtcct |      |
|            |            |            |            |            | cacatgetge |      |
|            |            |            |            |            | gaacgttacc |      |
|            |            |            |            |            | gaccgtgaaa |      |
|            |            |            |            |            | cagcaggaaa |      |
|            |            |            |            |            | tggttcgaca |      |
|            |            |            |            |            | ctgatcggtc |      |
|            |            |            |            |            | ccttctagtt |      |
|            |            |            |            |            | ggtgccactc |      |
|            |            |            |            |            | aggtgtcatt |      |
|            |            |            |            |            | gacaatagca |      |
|            |            |            |            |            | tgacccggtt |      |
|            |            |            |            |            | tgtccacgcc |      |
|            |            |            |            |            | gctccgcctt |      |
|            |            |            |            |            | ccaccaaacc |      |
|            |            |            |            |            | gtgcagaggg |      |
| -          | 5 5 5      |            |            |            |            |      |

```
agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat tttaaggcca 4080
tgatttaagg ccatcatggc cttaatcttc cgcttcctcg ctcactgact cgctgcgctc 4140
ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac 4200
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 4260
ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 4320
caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc 4380
gtttccccct ggaagctccc tegtgegete teetgttccq accetgeeqe ttaccqqata 4440
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 4500
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca 4560
gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga 4620
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4680
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4740
tatetgeget etgetgaage eagttacett eggaaaaaga gttggtaget ettgateegg 4800
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacqcqcaq 4860
aaaaaaagga teteaagaag ateetttgat ettttetaeg gggtetgaeg eteagtggaa 4920
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 4980
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5040
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 5100
atccatagtt gcctgactcg ggggggggg gcgctgaggt ctgcctcgtg aagaaggtgt 5160
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggaqccacq 5220
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5280
ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag caaaaqttcg 5340
atttattcaa caaagcegee gteeegteaa gteagegtaa tgetetgeea gtgttacaac 5400
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttattc 5460
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5520
teacegagge agttecatag gatggeaaga teetggtate ggtetgegat teegactegt 5580
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5640
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5700
actigiticaa caggicagee attacgeteg teatcaaaat cactegeate aaccaaaceg 5760
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 5820
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 5880
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg gatcgcagtg 5940
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6000
aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 6060
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6120
gcacctgatt gcccgacatt atcgcgagec catttatace catataaatc agcatecatg 6180
ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6240
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6300
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccattat 6360
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6420
aataaacaaa taggggttee gegeacattt eeeegaaaag tgeeacetga egtetaagaa 6480
accattatta tcatgacatt aacctataaa aataggegta teaegaggee ettteqte
```

```
<210> 146
<211> 671B
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 146
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
```

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 egteaatggg agtttgtttt ggeaceaaaa teaaegggae ttteeaaaat gtegtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920 ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040 gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100 acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160 cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220 ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280 aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340 ccaacgacat gaacaaagaa atccgtaact gctccttcgc atgcatcaac tgcaacgctt 2400 ccaccatcac ccaggettge ccgaaagtta acttcgacce gatcccgatc cactactgcg 2460 ctccggctgg ttacgctatc ctgaaatgca acaacaaaac cttctccggt aaaggtccgt 2520 gcaacaacgt ttccaccgtt cagtgcaccc atggtatcaa accggttgtt tccacccagc 2580 tgctgctgaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2640 acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcgtt tgcacccgtc 2700 cgaacaacaa cacccgtaaa tccatgcgta tcggtccggg tcagaccttc tacgctaccg 2760 gtgacatcat cggtgacatc cgtcaggctt actgcaacat ctccggttcc aaatggaacg 2820 aaaccctgaa acgtgttaaa gaaaaactgc aggaaaacta caacaacaac aaaaccatca 2880 aattegetee gteeteeggt ggtgaeetgg aaatcaceae ceaeteette aactgeegtg 2940 gtgaattett ctactgcaac accaccegte tgttcaacaa caacgctacc gaagacgaaa 3000 ccatcaccct gccgtgccgt atcaaacaga tcatcaacat gtggcagggt gttggtcgtg 3060 ctatgtacgc tccgccgatc gctggtaaca tcacctgcaa atccaacatc accggtctgc 3120 tgctggttcg tgacggtggt gaagacaaca aaaccgaaga aatcttccgt ccgggtggtg 3180 gtaacatgaa agacaactgg cgttccgaac tgtacaaata caaagttatc gaactgaaac 3240 cgctgggtat cgctccgacc ggtgctaagc ttaccgttca ggctcgtcag ctgctgtcct 3300 ccatcgttca gcagcagtcc aacctgctgc gtgctatcga agctcagcag cacatgctgc 3360

| agctgaccgt | ttggggtatc | aaacaqctqc | agacccgtgt | tetagetate | gaacgttacc | 3420 |
|------------|------------|------------|------------|------------|------------|------|
| tgaaagacca | gcagctcgag | atctqqaaca | acatgacctg | gatggaatgg | gaccgtgaaa | 3480 |
| tctccaacta | caccgacacc | atctaccotc | tactagaaga | choccagaco | Cancannaaa | 3540 |
| aaaacgaaaa | agacctgctg | actitagact | cctogaaaa  | cctataatca | taattaaaa  | 3500 |
| tctccaacto | gctgtggtac | atcaaaatct | tratratoat | cattagtagt | ctastcasta | 3660 |
| tocotateat | cttcgctgtt | ctotccatct | gaggatgag  | atetectete | ccgaccggtc | 3730 |
| gccagccatc | tgttgtttgc | ccctcccca  | taccttcctt | accepteges | agtagesete | 3720 |
| ccactotect | ttcctaataa | aatgaggaaa | ttacatcaca | ttototosat | ggtgccactc | 3780 |
| ctattctcc  | gaatagaata | aacgaggaaa | acasaggga  | cogcocyage | aggigicate | 3840 |
| aggetactag | gggtggggtg | gggtaggata | gcaaggggga | ggattgggaa | gacaatagca | 3900 |
| cetectors  | ggatgcggtg | ggetttatgg | gracecagge | gergaagaat | tgacccggtt | 3960 |
| catacttatt | cagaaagaag | caggeacate | ceettetetg | Lgacacaccc | tgtccacgcc | 4020 |
| dashacasa  | agttccagcc | ccacccacag | gacactcata | gercaggagg | gctccgcctt | 4080 |
| Caaceccace | cgctaaagta | cccggagegg | CCCCCCCCC  | cctcatcagc | ccaccaaacc | 4140 |
| adacctagee | tccaagagtg | ggaagaaact | aaagcaagat | aggctattaa | gtgcagaggg | 4200 |
| agagaaaacg | cctccaacat | grgaggaagt | aatgagagaa | atcatagaat | tttaaggcca | 4260 |
| cgatttaagg | ccatcatggc | cttaatcttc | cgcttcctcg | ctcactgact | cgctgcgctc | 4320 |
| ggtcgttcgg | ctgcggcgag | cggtatcagc | tcactcaaag | gcggtaatac | ggttatccac | 4380 |
| agaatcaggg | gataacgcag | gaaagaacat | gtgagcaaaa | ggccagcaaa | aggccaggaa | 4440 |
| ccgtaaaaag | gccgcgttgc | tggcgttttt | ccataggctc | cgcccccctg | acgagcatca | 4500 |
| caaaaatcga | cgctcaagtc | agaggtggcg | aaacccgaca | ggactataaa | gataccaggc | 4560 |
| grraceet   | ggaagctccc | tegtgegete | tcctgttccg | accctgccgc | ttaccggata | 4620 |
| cctgtccgcc | tttctccctt | cgggaagcgt | ggcgctttct | catageteae | gctgtaggta | 4680 |
| tctcagttcg | gtgtaggtcg | ttcgctccaa | gctgggctgt | gtgcacgaac | ccccgttca  | 4740 |
| gcccgaccgc | tgcgccttat | ccggtaacta | tcgtcttgag | tccaacccgg | taagacacga | 4800 |
| cttatcgcca | ctggcagcag | ccactggtaa | caggattagc | agagcgaggt | atgtaggcgg | 4860 |
| tgctacagag | ttcttgaagt | ggtggcctaa | ctacggctac | actagaagaa | cagtatttgg | 4920 |
| tatetgeget | ctgctgaagc | cagttacctt | cggaaaaaga | gttggtagct | cttgatccgg | 4980 |
| caaacaaacc | accgctggta | gcggtggttt | ttttgtttgc | aagcagcaga | ttacgcgcag | 5040 |
| aaaaaaagga | tctcaagaag | atcetttgat | cttttctacg | gggtctgacg | ctcagtggaa | 5100 |
| cgaaaactca | cgttaaggga | ttttggtcat | gagattatca | aaaaggatct | tcacctagat | 5160 |
| ccttttaaat | taaaaatgaa | gttttaaatc | aatctaaagt | atatatgagt | aaacttggtc | 5220 |
| tgacagttac | caatgcttaa | tcagtgaggc | acctatctca | gcgatctgtc | tatttcgttc | 5280 |
| acccataget | gcctgactcg | aaaaaaaaa  | gcgctgaggt | ctgcctcgtg | aagaaggtgt | 5340 |
| tgctgactca | taccaggeet | gaatcgcccc | atcatccagc | cagaaagtga | gggagccacg | 5400 |
| grrgargaga | gctttgttgt | aggtggacca | gttggtgatt | ttgaactttt | gctttgccac | 5460 |
| ggaacggtct | gcgttgtcgg | gaagatgcgt | gatctgatcc | ttcaactcag | caaaagttcg | 5520 |
| acttattcaa | caaagccgcc | gtcccgtcaa | gtcagcgtaa | tgctctgcca | gtgttacaac | 5580 |
| caattaacca | attctgatta | gaaaaactca | tcgagcatca | aatgaaactg | caatttattc | 5640 |
| acaccaggat | tatcaatacc | atattttga  | aaaagccgtt | tctgtaatga | aggagaaaac | 5700 |
| tcaccgagge | agttccatag | gatggcaaga | tcctggtatc | ggtctgcgat | tccgactcgt | 5760 |
| ccaacatcaa | tacaacctat | taatttcccc | tcgtcaaaaa | taaggttatc | aagtgagaaa | 5820 |
| tcaccatgag | tgacgactga | atccggtgag | aatggcaaaa | gcttatgcat | ttctttccag | 5880 |
| acttgttcaa | caggccagcc | attacgctcg | tcatcaaaat | cactcgcatc | aaccaaaccg | 5940 |
| ttattcattc | gtgattgcgc | ctgagcgaga | cgaaatacgc | gatcgctgtt | aaaaggacaa | 6000 |
| ttacaaacag | gaatcgaatg | caaccggcgc | aggaacactg | ccagcgcatc | aacaatattt | 6060 |
| tcacctgaat | caggatattc | ttctaatacc | tggaatgctg | ttttcccggg | gatcgcagtg | 6120 |
| gtgagtaacc | atgcatcatc | aggagtacgg | ataaaatgct | tgatggtcgg | aagaggcata | 6180 |
| aattccgtca | gccagtttag | tctgaccatc | tcatctgtaa | catcattggc | aacgctacct | 6240 |
| ttgccatgtt | tcagaaacaa | ctctggcgca | tegggettee | catacaatcg | atagattgtc | 6300 |
| gcacctgatt | gcccgacatt | atcgcgagcc | catttatacc | catataaatc | agcatccatg | 6360 |
| ccggaattta | atcgcggcct | cgagcaagac | gtttcccgtt | gaatatggct | cataacaccc | 6420 |
| cctgtattac | tgtttatgta | agcagacagt | tttattgttc | atgatgatat | atttttatct | 6480 |
| rgrgcaatgt | aacatcagag | attttgagac | acaacgtggc | tttcccccc  | ccccattat  | 6540 |
| cgaagcattt | atcagggtta | ttgtctcatg | agcggataca | tatttgaatg | tatttagaaa | 6600 |
| aacaacaaa  | taggggttcc | gcgcacattt | ccccgaaaag | tgccacctga | cgtctaagaa | 6660 |
|            |            |            |            |            |            |      |

6718

accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc

```
<210> 147
<211> 6637
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 147
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccqccatqtt qacattqatt attqactaqt tattaataqt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
cataqtaacq ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccqcccca ttqacqcaaa tqqqcqqtaq gcqtqtacqq tqqqaqqtct atataaqcaq 900
agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac egggaeegat eeageeteeg eggeegggaa eggtgeattg gaaegeggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080
tettatgeat getatactgt tittggettg gggeetatae acceeegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acqaccqtqa aqttcacaac qtttgqgcta cccacqcttg cgttccgacc gacccgaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340
ccaacgacat gaacaaagaa atccgtaact gctccttcgc atgcatcaac tgcaacgctt 2400
ccaccateac ccaggettge ccgaaagtta acttegacce gatecegate cactactgeg 2460
ctccqqctqq ttacgctatc ctgaaatgca acaacaaaac cttctccggt aaaggtccgt 2520
```

|            |            |            |            |            | tccacccage |      |
|------------|------------|------------|------------|------------|------------|------|
| tgctgctgaa | cggttccctg | gctgaaaaag | aaatcatcat | ccgttccgaa | aacctgaccg | 2640 |
| acaacgttaa | aaccatcatc | gttcacctga | acaaatccgt | tgaaatcgtt | tgcacccgtg | 2700 |
|            |            |            |            |            | cgtgttaaag |      |
|            |            |            |            |            | tcctccggtg |      |
|            |            |            |            |            | tactgcaaca |      |
| ccacccgtct | gttcaacaac | aacgctaccg | aagacgaaac | catcaccctg | ccgtgccgta | 2940 |
| tcaaacagat | catcaacatg | tggcagggtg | ttggtcgtgc | tatgtacgct | ecgccgatcg | 3000 |
|            |            |            |            |            | gacggtggtg |      |
|            |            |            |            |            | gacaactggc |      |
|            |            |            |            |            | gctccgaccg |      |
|            |            |            |            |            | cagcagtcca |      |
| acctgctgcg | tgctatcgaa | gctcagcagc | acatgctgca | gctgaccgtt | tggggtatca | 3300 |
|            |            |            |            |            | cagctcgaga |      |
|            |            |            |            |            | accgacacca |      |
| tctaccgtct | gctggaagac | tcccagaccc | agcaggaaaa | aaacgaaaaa | gacctgctgg | 3480 |
|            |            |            |            |            | ctgtggtaca |      |
| tcaaaatctt | catcatgatc | gttggtggtc | tgatcggtct | gcgtatcatc | ttcgctgttc | 3600 |
| tgtccatctg | aggatccaga | tctgctgtgc | cttctagttg | ccagccatct | gttgtttgcc | 3660 |
| cctcccccgt | gccttccttg | accctggaag | gtgccactcc | cactgtcctt | tectaataaa | 3720 |
| atgaggaaat | tgcatcgcat | tgtctgagta | ggtgtcattc | tattctgggg | ggtggggtgg | 3780 |
| ggcaggacag | caagggggag | gattgggaag | acaatagcag | gcatgctggg | gatgcggtgg | 3840 |
| gctctatggg | tacccaggtg | ctgaagaatt | gacccggttc | ctcctgggcc | agaaagaagc | 3900 |
|            |            |            |            |            | gttccagccc |      |
|            |            |            |            |            | gctaaagtac |      |
|            |            |            |            |            | ccaagagtgg |      |
|            |            |            |            |            | ctccaacatg |      |
|            |            |            |            |            | catcatggcc |      |
|            |            |            |            |            | tgcggcgagc |      |
|            |            |            |            |            | ataacgcagg |      |
|            |            |            |            |            | ccgcgttgct |      |
| ggcgttttc  | cataggetee | gccccctga  | cgagcatcac | aaaaatcgac | gctcaagtca | 4440 |
|            |            |            |            |            | gaagctccct |      |
|            |            |            |            |            | ttctcccttc |      |
|            |            |            |            |            | tgtaggtcgt |      |
|            |            |            |            |            | gcgccttatc |      |
|            |            |            |            |            | tggcagcagc |      |
|            |            |            |            |            | tcttgaagtg |      |
|            |            |            |            |            | tgctgaagcc |      |
| agttaccttc | ggaaaaagag | ttggtagctc | ttgatccggc | aaacaaacca | ccgctggtag | 4920 |
|            |            |            |            |            | ctcaagaaga |      |
|            |            |            |            |            | gttaagggat |      |
|            |            |            |            |            | aaaaatgaag |      |
|            |            |            |            |            | aatgcttaat |      |
|            |            |            |            |            | cctgactcgg |      |
|            |            |            |            |            | accaggectg |      |
|            |            |            |            |            | ctttgttgta |      |
|            |            |            |            |            | cgttgtcggg |      |
|            |            |            |            |            | aaagccgccg |      |
|            |            |            |            |            | ttctgattag |      |
|            |            |            |            |            | atcaatacca |      |
|            |            |            |            |            | gttccatagg |      |
|            |            |            |            |            | acaacctatt |      |
|            |            |            |            |            | gacgactgaa |      |
|            |            |            |            |            | aggccagcca |      |
|            |            |            |            |            | 55 -5554   |      |

```
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 5880 tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 5940 aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 6000 tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tgcatcatca 6060 ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6120 ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6180 tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6240 tcgcgagccc atttataccc atataggca acaaccacc ttggaatttaa tcgcgggcctc 6300 gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6360 gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6420 tcttggacac caacgtggct ttccccccc ccccattatt gaagcattta tcagggttat 6480 tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6540 cccgacatttc cccgaaaagt cccgaggccc ttcccccc ttcgtcc cattatta catgacatta 6600 acctataaaa ataggcgtat cacgaggccc ttccgtc
```

<210> 148
<211> 6598
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

tegegegettt eggegatgae ggegaaaace teegacacat geageteeeg gagaeggea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cqqtaaatqq 420 cccgcctggc tgaccgccca acgaccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccq cqqccgggaa cqqtqcattq qaacqcqqat 1020 teccegtgcc aagagtgacg taagtacege ctatagacte tataggcaca eccetttggc 1080 tettatgeat getatactgt ttttggettg gggeetatac acceegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac qtqttccqqa 1440 catgggetet teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 aggggctcat ggtcgctcgg cagetccttg ctcctaacag tggaggccag acttaggcac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaqqcaqcq 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

| attacaatac  | tattaacaat  | ggagggagt  | ataatetaaa  | cagtactogt | tactaccaca  | 1900 |
|-------------|-------------|------------|-------------|------------|-------------|------|
| cacacca     | gacataatag  | ctgacagact | aacagactgt  | tcctttccat | gggtcttttc  | 1860 |
| tocaotcacc  | gtcgtcgaca  | catataatca | gatatogogo  | ccgctctaga | gagaccccc   | 1920 |
| ccatgcgtgt  | tcataatata  | ccacataact | aaccacaata  | gtggatgtgg | gatategtea  | 1900 |
| gtttctggat  | gatcatcatc  | taccatatta | ttootaacat  | gtgggttacc | gtttactaca  | 2040 |
| gtattccat   | ttggaccgac  | gctaaaacca | ccctattcta  | cgcttccgac | accasages   | 2100 |
| acgaccgtga  | agttcacaac  | gettaaacta | cccacactta  | cgttccgacc | accaaagccc  | 2100 |
| cocaggaaat. | cattetagaa  | aacottacco | aaaacttcaa  | catgtggaaa | accegaacc   | 2220 |
| traaccapat  | acacasaas   | atcatctccc | tataaccccaa | gtccctgaaa | aacgacatgg  | 2220 |
| aactgacccc  | actatacatt  | accetocact | acaccaacac  | taccttcaaa | aacaacetta  | 2200 |
| ccaacgacat  | daacaaadaa  | atccctaact | getecttege  | atgcatcaac | tacaacytta  | 2340 |
| ccaccatcac  | ccaggettge  | ccassatts  | acttccacce  | gatcccgatc | cycaacyctt  | 2400 |
| ctccactaa   | thacachata  | chassatacs | acettgatt   | cttctccggt | caccaccgcg  | 2460 |
| acaacaacat  | ttccaccatt  | cagaaacgca | ataataaaac  | accepte    | taaaggteege | 2520 |
| tactactase  | coattacta   | agtassassa | acggcaccaa  | accggttgtt | ccaccage    | 2580 |
| acaacottaa  | aaccatcatc  | gttgaaaaag | adaccaccat  | ccgttccgaa | aacctgaccg  | 2640 |
| ctaccectta  | chaccaccatc | tacathan   | acaaacccgc  | tgaaatcgtt | tgcacccgtg  | 2700 |
| aaaaactaaa  | ggaaaacacc  | ceeggeeea  | aatggaacga  | aaccctgaaa | cgcgccaaag  | 2760 |
| atazastas   | ggaaaactac  | aacaacaaca | aaaccatcaa  | attcgctccg | tcctccggtg  | 2820 |
|             |             |            |             | tgaattcttc |             |      |
| cegetageet  | geegtgeegt  | accaaacaga | ccaccacat   | gtggcagggt | grtggtcgtg  | 2940 |
| tratgtacge  | teegeegate  | gctggtaaca | tcacctgcaa  | atccaacatc | accggtctgc  | 3000 |
| tgetggtteg  | tgacggtggt  | gaagacaaca | aaaccgaaga  | aatcttccgt | ccgggtggtg  | 3060 |
| graacargaa  | agacaactgg  | cgttccgaac | tgtacaaata  | caaagttatc | gaactgaaac  | 3120 |
| cgctgggtat  | cgctccgacc  | ggtgctaagc | ttaccgttca  | ggctcgtcag | ctgctgtcct  | 3180 |
| ccatcgttca  | gcagcagtcc  | aacctgctgc | gtgctatcga  | agctcagcag | cacatgctgc  | 3240 |
| agctgaccgt  | ttggggtatc  | aaacagctgc | agacccgtgt  | tctggctatc | gaacgttacc  | 3300 |
| tgaaagacca  | gcagctcgag  | atctggaaca | acatgacctg  | gatggaatgg | gaccgtgaaa  | 3360 |
| tctccaacta  | caccgacacc  | atctaccgtc | tgctggaaga  | ctcccagacc | cagcaggaaa  | 3420 |
| aaaacgaaaa  | agacctgctg  | gctctggact | cctggaaaaa  | cctgtggtcc | tggttcgaca  | 3480 |
| tetecaactg  | gctgtggtac  | atcaaaatct | tcatcatgat  | cgttggtggt | ctgatcggtc  | 3540 |
| tgcgtatcat  | cttcgctgtt  | ctgtccatct | gaggatccag  | atctgctgtg | ccttctagtt  | 3600 |
| gccagccatc  | tgttgtttgc  | ccctccccg  | tgccttcctt  | gaccctggaa | ggtgccactc  | 3660 |
| ccactgtcct  | ttcctaataa  | aatgaggaaa | ttgcatcgca  | ttgtctgagt | aggtgtcatt  | 3720 |
| ctattctggg  | gggtggggtg  | gggcaggaca | gcaaggggga  | ggattgggaa | gacaatagca  | 3780 |
| ggcatgctgg  | ggatgcggtg  | ggctctatgg | gtacccaggt  | gctgaagaat | tgacccggtt  | 3840 |
|             |             |            |             | tgacacaccc |             |      |
| cctggttctt  | agttccagcc  | ccactcatag | gacactcata  | gctcaggagg | gctccgcctt  | 3960 |
| caatcccacc  | cgctaaagta  | cttggagcgg | tetetecete  | cctcatcagc | ccaccaaacc  | 4020 |
| aaacctagcc  | tccaagagtg  | ggaagaaatt | aaagcaagat  | aggctattaa | gtgcagaggg  | 4080 |
| agagaaaatg  | cctccaacat  | gtgaggaagt | aatgagagaa  | atcatagaat | tttaaggcca  | 4140 |
| tgatttaagg  | ccatcatggc  | cttaatcttc | cgcttcctcg  | ctcactgact | cgctgcgctc  | 4200 |
| ggtcgttcgg  | ctgcggcgag  | cggtatcagc | tcactcaaag  | gcggtaatac | ggttatccac  | 4260 |
| agaatcaggg  | gataacgcag  | gaaagaacat | gtgagcaaaa  | ggccagcaaa | aggccaggaa  | 4320 |
| ccgtaaaaag  | gccgcgttgc  | tggcgttttt | ccataggctc  | cgccccctg  | acgagcatca  | 4380 |
| caaaaatcga  | cgctcaagtc  | agaggtggcg | aaacccgaca  | ggactataaa | gataccaggc  | 4440 |
|             |             |            |             | accetgeege |             |      |
| cctgtccgcc  | tttctccctt  | cgggaagcgt | ggcgctttct  | catageteae | gctgtaggta  | 4560 |
| tctcagttcg  | gtgtaggtcg  | ttcgctccaa | gctgggctgt  | gtgcacgaac | ccccgttca   | 4620 |
| gcccgaccgc  | tgcgccttat  | ccggtaacta | tcgtcttgag  | tccaacccgg | taagacacga  | 4680 |
| cttatcgcca  | ctggcagcag  | ccactggtaa | caggattagc  | agagcgaggt | atgtaggcgg  | 4740 |
| tgctacagag  | ttcttgaagt  | ggtggcctaa | ctacggctac  | actagaagaa | cagtatttgg  | 4800 |
| tatctgcgct  | ctgctgaagc  | cagttacctt | cggaaaaaga  | gttggtagct | cttgatccgg  | 4860 |
| caaacaaacc  | accgctggta  | gcggtggttt | ttttgtttgc  | aagcagcaga | ttacgcgcag  | 4920 |
| aaaaaagga   | tctcaagaag  | atcctttgat | cttttctacg  | gggtctgacg | ctcagtggaa  | 4980 |
| cgaaaactca  | cgttaaggga  | ttttggtcat | gagattatca  | aaaaggatct | tcacctagat  | 5040 |

```
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5100
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 5160
atccatagtt gcctgactcg ggggggggg gcgctgaggt ctgcctcgtg aagaaggtgt 5220
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacq 5280
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5340
ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag caaaagttcg 5400
atttattcaa caaageegee gteeegteaa gteagegtaa tgetetgeea gtgttacaac 5460
caattaacca attetgatta gaaaaactca tegageatca aatgaaactg caatttatte 5520
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5580
tcaccgagge agttccatag gatggcaaga tcctggtatc ggtctgcgat tccgactcgt 5640
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5700
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5760
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 5820
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 5880
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 5940
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg gatcgcagtg 6000
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6060
aatteegtea geeagtttag tetgaceate teatetgtaa cateattgge aacgetacet 6120
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg ataqattqtc 6180
gracetgatt greegacatt ategegagee catttatace catataaate ageatecatg 6240
ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6300
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6360
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttcccccc ccccattat 6420
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6480
aataaacaaa taggggttee gegeacattt eeeegaaaag tgecacetga egtetaagaa 6540
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcqtc 6598
```

<212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tegegegetet eggegatgae ggegaaaace tetgaeacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020

<210> 149 <211> 6679

| tacacataca | 224244444   | +ab         |            | <b>L</b> _ <b>L</b> |            |      |
|------------|-------------|-------------|------------|---------------------|------------|------|
| tettegegee | aagagugacg  | Laagtaccgc  | ctatagacte | tataggcaca          | ccccttggc  | 1080 |
| Lectatgeat | getataetgt  | cccggcttg   | gggcctatac | acccccgctt          | ccttatgcta | 1140 |
| Laggigatgg | tatagettag  | cctataggtg  | tgggttattg | accattattg          | accactcccc | 1200 |
| tattggtgac | gatacttcc   | attactaatc  | cataacatgg | ctctttgcca          | caactatctc | 1260 |
| tattggctat | atgccaatac  | tetgteette  | agagactgac | acggactctg          | tatttttaca | 1320 |
| ggatggggtc | ccatttatta  | tttacaaatt  | cacatataca | acaacgccgt          | cccccgtgcc | 1380 |
| cgcagttttt | attaaacata  | gcgtgggatc  | tccacgcgaa | tctcgggtac          | gtgttccgga | 1440 |
| catgggctct | tctccggtag  | cggcggagct  | tccacatccg | agccctggtc          | ccatgcctcc | 1500 |
| agcggctcat | ggtcgctcgg  | cagctccttg  | ctcctaacag | tggaggccag          | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac  | cagtgtgccg  | cacaaggccg | tggcggtagg          | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg  | ggctcgcacg  | gctgacgcag | atggaagact          | taaggcagcg | 1680 |
| gcagaagaag | atgcaggcag  | ctgagttgtt  | gtattctgat | aagagtcaga          | ggtaactccc | 1740 |
| gttgcggtgc | tgttaacggt  | ggagggcagt  | gtagtctgag | cagtactcgt          | tgctgccqcq | 1800 |
|            |             |             |            | tcctttccat          |            |      |
| tgcagtcacc | gtcgtcgaca  | cgtgtgatca  | gatatcgcgg | ccgctctaga          | gatatogoca | 1920 |
|            |             |             |            | gtggatgtgg          |            |      |
|            |             |             |            | gtgggttacc          |            |      |
| gtgttccggt | ttggaccgac  | gctaaaacca  | ccctqttctq | cgcttccgac          | accaaagcct | 2100 |
| acgaccgtga | agttcacaac  | gtttgggcta  | cccacactta | cgttccgacc          | gacccgaacc | 2160 |
| cqcaqqaaat | cattetagaa  | aacgttaccg  | aaaacttcaa | catgtggaaa          | aacgacatgg | 2220 |
| ttgaccagat | gcacgaagac  | atcatctccc  | tataggacca | gtccctgaaa          | ccatacatta | 2280 |
| aactgacccc | actatacatt  | accetocact  | gcaccaacgc | taccttcaaa          | aacaacutta | 2340 |
| ccaacgacat | gaacaaagaa  | atccotaact  | actecttede | atgcatcaac          | tacaacactt | 2400 |
| ccaccatcac | ccagacttac  | ccdaaadtta  | acttccaccc | gatcccgatc          | cachachaca | 2400 |
| ctccacta   | ttacgctatc  | ctgaaagcca  | accecgacec | cttctccggt          | caccaccgeg | 2400 |
| graacaacgt | ttccaccatt  | cagaaacyca  | ataatataa  | acceptes            | aaaggteegt | 2520 |
| tactactass | cactteesta  | gatassassa  | arggrateat | accggttgtt          | tecacecage | 2580 |
| acaacattaa | aaccatcatc  | gttgaaaaag  | adaccaccat | ccgttccgaa          | aacctgaccg | 2640 |
| ccaacgecaa | caccattana  | teesteeste  | tagatagaga | tgaaatcgtt          | tgeaceegte | 2700 |
| atascatcat | cactegeaaa  | cataragests | actacaacat | tcagaccttc          | Lacyctaccy | 2760 |
| aaaccctcac | acatattaaa  | gecaggeee   | accycaacac | ctccggttcc          | aaatggaacg | 2820 |
| aaaccccgaa | atectecast  | gaaaaactgc  | aggaaaacta | caacaacaac          | aaaaccatca | 2880 |
| atasattatt | chachecase  | agegateegg  | haacataaa  | ccactccttc          | aactgccgtg | 2940 |
| tataacaaaa | tattaataat  | actictates  | cyccycyccy | tatcaaacag          | atcatcaaca | 3000 |
| aatccaacat | caccagteta  | geratgrate  | cteegeegat | cgctggtaac          | atcacctgca | 3060 |
| aacccaacac | tacceggeeeg | cegeeggeee  | grgacggrgg | tgaagacaac          | aaaaccgaag | 3120 |
| agazzattat | cccgggcggc  | ggcaacacya  | tagatagetg | gcgttccgaa          | ctgtacaaat | 3180 |
| acadagetae | actactataa  | tesatesta   | regereegae | cggtgctaag          | cttaccgttc | 3240 |
| aggetegeea | gergergere  | cocatogue   | ageageagee | caacctgctg          | egtgetateg | 3300 |
| thetecates | gcacacgccg  | cagetgaceg  | cccggggcac | caaacagctg          | cagacccgcg | 3360 |
| gastgassta | agaacgttac  | ctgaaagacc  | ageagetega | gatctggaac          | aacatgacct | 3420 |
|            |             |             |            | catctaccgt          |            |      |
| acteceagae | ccagcaggaa  | addadcgaaa  | aagacctgct | ggctctggac          | ccctggaaaa | 3540 |
| acctgtggte | ccggcccgae  | acccccaacc  | ggctgtggta | catcaaaatc          | ttcatcatga | 3600 |
| cegetggtgg | tetgateggt  | ctgcgtatca  | cctcgctgt  | tctgtccatc          | tgaggatcca | 3660 |
| gatetgetgt | geettetagt  | tgccagccat  | ctgttgtttg | cccatcccc           | gtgccttcct | 3720 |
| cgaccccgga | aggrgccact  | cccactgtcc  | tttcctaata | aaatgaggaa          | attgcatcgc | 3780 |
| accgcctgag | taggtgtcat  | tctattctgg  | agaaraaaar | ggggcaggac          | agcaaggggg | 3840 |
| aggactggga | agacaatagc  | aggcatgctg  | gggatgcggt | gggctctatg          | ggtacccagg | 3900 |
| tgctgaagaa | rrgacccggt  | tectectggg  | ccagaaagaa | gcaggcacat          | ccccttctct | 3960 |
| grgacacacc | ctgtccacgc  | ccctggttct  | tagttccagc | cccactcata          | ggacactcat | 4020 |
| agcccaggag | ggctccgcct  | tcaatcccac  | ccgctaaagt | acttggagcg          | gtctctccct | 4080 |
| ccctcatcag | cccaccaaac  | caaacctagc  | ctccaagagt | gggaagaaat          | taaagcaaga | 4140 |
| taggctatta | agtgcagagg  | gagagaaaat  | gcctccaaca | tgtgaggaag          | taatgagaga | 4200 |
| aatcatagaa | ttttaaggcc  | atgatttaag  | gccatcatgg | ccttaatctt          | ccgcttcctc | 4260 |
| gctcactgac | tegetgeget  | cggtcgttcg  | gctgcggcga | gcggtatcag          | ctcactcaaa | 4320 |
|            |             |             |            |                     |            |      |

```
ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 4380
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct 4440
ecgececet gacgageate acaaaaateg acgeteaagt cagaggtgge gaaaceegae 4500
aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc 4560
gaccetgeeg ettaceggat acetgteege ettteteeet tegggaageg tggegettte 4620
tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg 4680
tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga 4740
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag 4800
cagagegagg tatgtaggeg gtgctacaga gttcttgaag tggtggccta actacggcta 4860
cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag 4920
agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg 4980
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 5040
ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc 5100
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag 5160
tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 5220
tetgeetegt gaagaaggtg ttgetgacte ataccaggee tgaategeee cateatecag 5340
ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc agttggtgat 5400
cttcaactca gcaaaagttc gatttattca acaaagccgc cgtcccgtca agtcagcgta 5520
atgetetgee agtgttacaa ecaattaace aattetgatt agaaaaacte ategageate 5580
aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5640
ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag atcctggtat 5700
eggtetgega tteegacteg teeaacatea atacaaceta ttaattteee etegteaaaa 5760
ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga gaatggcaaa 5820'
agettatgca tttettteca gaettgttea acaggeeage cattacgete gteatcaaaa 5880
tcactcgcat caaccaaacc gttattcatt cgtgattgcg cctgagcgag acgaaatacg 5940
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact 6000
gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6060
gttttcccgg ggatcgcagt ggtgagtaac catgcatcat caggagtacg gataaaatgc 6120
ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat ctcatctgta 6180
acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc 6240
ccatacaatc gatagattgt cgcacctgat tgcccgacat tatcgcgagc ccatttatac 6300
ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga cgtttcccgt 6360
tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag ttttattgtt 6420
catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6480
ctttcccccc cccccatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 6540
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 6600
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6660
atcacgaggc cctttcqtc
                                                              6679
```

```
<210> 150
<211> 6748
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVRl012x/s containing HTV genes
<400> 150
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggggg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
```

|            |            |            |            | •           |            |       |
|------------|------------|------------|------------|-------------|------------|-------|
| accatatgcg | gtgtgaaata | ccgcacagat | gcgtaaggag | aaaataccgc  | atcagattgg | 240   |
| ctattggcca | ttgcatacgt | tgtatccata | tcataatatg | tacatttata  | ttggctcatg | 300   |
| tccaacatta | ccgccatgtt | gacattgatt | attgactagt | tattaatagt  | aatcaattac | 360   |
| ggggtcatta | gttcatagcc | catatatgga | gttccgcgtt | acataactta  | cggtaaatgg | 420   |
| cccgcctggc | tgaccgccca | acgacccccg | cccattgacg | tcaataatga  | cgtatgttcc | 480   |
|            |            |            | acgtcaatgg |             |            |       |
| tgcccacttg | gcagtacatc | aagtgtatca | tatgccaagt | acgcccccta  | ttgacgtcaa | 600   |
| tgacggtaaa | tggcccgcct | ggcattatgc | ccagtacatg | accttatggg  | actttcctac | 660   |
| ttggcagtac | atctacgtat | tagtcatcgc | tattaccatg | gtgatgcggt  | tttggcagta | 720   |
| catcaatggg | cgtggatagc | ggtttgactc | acggggattt | ccaagtctcc  | accccattga | 780   |
| cgtcaatggg | agtttgttt  | ggcaccaaaa | tcaacgggac | tttccaaaat  | gtcgtaacaa | 840   |
| ctccgcccca | ttgacgcaaa | tgggcggtag | gcgtgtacgg | tgggaggtct  | atataagcag | 900   |
| agctcgttta | gtgaaccgtc | agatcgcctg | gagacgccat | ccacgctgtt  | ttgacctcca | 960   |
| tagaagacac | cgggaccgat | ccagcctccg | cggccgggaa | cggtgcattg  | gaacgcqqat | 1020  |
| teccegtgee | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca  | cccctttggc | 1080  |
| tcttatgcat | gctatactgt | ttttggcttg | gggcctatac | acccccgctt  | ccttatgcta | 1140  |
| taggtgatgg | tatagcttag | cctataggtg | tgggttattg | accattattg  | accactccc  | 1200  |
| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttgcca  | caactatctc | 1260  |
| tattggctat | atgccaatac | tctgtccttc | agagactgac | acqqactctq  | tatttttaca | 1320  |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacqccqt  | ccccatacc  | 1380  |
| cgcagttttt | attaaacata | qcqtqqqatc | tccacgcgaa | tctcgggtac  | gtgttccgga | 1440  |
| catgggctct | teteeggtag | cadcadadct | tccacatccg | agccctggtc  | ccatgcctcc | 1500  |
| agcggctcat | gatcactcaa | cageteettg | ctcctaacag | tagaagaccag | achtaggcac | 1560  |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | taacaataaa  | atatatatat | 1620  |
| gaaaatgagc | gtggagattg | aactcacaca | gctgacgcag | atogaagact  | taagggggg  | 1690  |
| gcagaagaag | atgcaggcag | ctgagttgtt | gtattctgat | aagagtcaga  | ggtaactccc | 1740  |
| attacaatac | tottaacoot | ggagggagt. | gtagtctgag | cagtactcgt  | tactaccaca | 1800  |
| cacaccacca | gacataatag | ctgacagact | aacagactgt | teettteeat  | agatettte  | 1960  |
| tgcagtcacc | gtcgtcgaca | catataatca | gatatcgcgg | ccactctaga  | gatatogoca | 1000  |
| ccatgcgtgt | tcataatatc | ccacataact | ggccgcagtg | ataatataa   | gatategtea | 1000  |
| gtttctggat | gatcatcatc | taccatatta | ttggtaacat | gtggatgtgg  | gtttactaca | 2040  |
| gtattccaat | ttagaccaac | ochasascca | ccctgttctg | gegggeeace  | accease    | 23.00 |
| acgaccgtga | agttcacaac | atttagacta | cccacgcttg | cattecase   | accaaageet | 2100  |
| cacaaaaaa  | cattetagaa | aacattacca | aaaacttcaa | gatatagaaa  | gaccegaace | 2190  |
| ttgaccagat | acacasaas  | atcatctccc | tgtgggacca | cacgeggaaa  | aacyacatyy | 2220  |
| aactgacccc | actatacatt | accatecece | gcaccaacgc | tacetteaaa  | cogracatta | 2280  |
| ccaacgacat | gaacaaagaa | atcooteact | gctccttcaa | cacciccaaa  | aacaacytta | 2340  |
| acaaaaaaca | gaacaaagaa | actogeaact | accgtccgga | cattactact  | gaaacccgcg | 2400  |
| accotaacaa | ctccaacaac | tocastsos  | testastas  | categetetg  | tegaaagaaa | 2460  |
| cccaaactta | cccaacaac  | aacttcaaca | tcctgatcaa | ccgcaacgct  | cccaccatca | 2520  |
| attacactat | catananta  | aaccccgacc | cgatcccgat | tasses      | gereeggerg | 2580  |
| tttccaccat | teaghacasa | antantana  | ccttctccgg | taaaggteeg  | cgcaacaacg | 2640  |
| accetteest | ggetgeace  | catggtatea | aaccggttgt | cccacccag   | ctgctgctga | 2700  |
| acggeeeee  | ggctgaaaaa | gaaatcatca | tccgttccga | aaacctgacc  | gacaacgtta | 2760  |
| adaccaccac | chaccatte  | aacaaacccg | ttgaaatcgt | ctgcacccgt  | gctagcgctt | 2820  |
| actgcaacat | ccccggttcc | aaacggaacg | aaaccctgaa | acgtgttaaa  | gaaaaactgc | 2880  |
| aggadaacta | caacaacaac | aaaaccacca | aattcgctcc | gtcctccggt  | ggtgacctgg | 2940  |
| habbanana  | ccactccttc | aactgccgtg | gtgaattctt | ctactgcaac  | accacccgtc | 3000  |
| tgtttaacaa | caacgctacc | gaagacgaaa | ccatcaccct | gccgtgccgt  | atcaaacaga | 3060  |
| teaccaacat | gcggcagggt | gctggtcgtg | ctatgtacgc | cccgccgatc  | gctggtaaca | 3120  |
| ccacctgcaa | acccaacatc | accggtctgc | tgctggttcg | rgacggtggt  | gaagacaaca | 3180  |
| aaaccgaaga | aatcttccgt | ccgggtggtg | gtaacatgaa | agacaactgg  | cgttccgaac | 3240  |
| tgtacaaata | caaagttatc | gaactgaaac | cgctgggtat | cgctccgacc  | ggtgctaagc | 3300  |
| ttaccgttca | ggctcgtcag | ctgctgtcct | ccatcgttca | gcagcagtcc  | aacctgctgc | 3360  |
| gtgctatcga | agcccagcag | cacatgctgc | agctgaccgt | ttggggtatc  | aaacagctgc | 3420  |
| agacccgtgt | tctggctatc | gaacgttacc | tgaaagacca | gcagctcgag  | atctggaaca | 3480  |

| acatgacctg | gatggaatgg | gaccgtgaaa | totocaacta | caccdacacc | atctaccotc | 3540 |
|------------|------------|------------|------------|------------|------------|------|
| tactagaaga | ctcccagacc | cagcaggaaa | aaaaccaaaa | agacctgctg | actictages | 3600 |
|            |            | tggttcgaca |            |            |            |      |
|            |            | ctgatcggtc |            |            |            |      |
|            |            |            |            |            |            |      |
| bassbaseb  | accegacgeg | ccttctagtt | gccagccacc | therese    | ccccccccg  | 3/80 |
|            |            | ggtgccactc |            |            |            |      |
|            |            | aggtgtcatt |            |            |            |      |
|            |            | gacaatagca |            |            |            |      |
|            |            | tgacccggtt |            |            |            |      |
|            |            | tgtccacgcc |            |            |            |      |
|            |            | gctccgcctt |            |            |            |      |
| tctctccctc | cctcatcagc | ccaccaaacc | aaacctagcc | tccaagagtg | ggaagaaatt | 4200 |
| aaagcaagat | aggctattaa | gtgcagaggg | agagaaaatg | cctccaacat | gtgaggaagt | 4260 |
| aatgagagaa | atcatagaat | tttaaggcca | tgatttaagg | ccatcatggc | cttaatcttc | 4320 |
|            |            | cgctgcgctc |            |            |            |      |
|            |            | ggttatccac |            |            |            |      |
|            |            | aggccaggaa |            |            |            |      |
|            |            | acgagcatca |            |            |            |      |
|            |            | gataccagge |            |            |            |      |
|            |            | ttaccggata |            |            |            |      |
|            |            | gctgtaggta |            |            |            |      |
|            |            |            |            |            |            |      |
|            |            | cccccgttca |            |            |            |      |
|            |            | taagacacga |            |            |            |      |
|            |            | atgtaggcgg |            |            |            |      |
|            |            | cagtatttgg |            |            |            |      |
|            |            | cttgatccgg |            |            |            |      |
|            |            | ttacgcgcag |            |            |            |      |
|            |            | ctcagtggaa |            |            |            |      |
| gagattatca | aaaaggatct | tcacctagat | ccttttaaat | taaaaatgaa | gttttaaatc | 5220 |
|            |            | aaacttggtc |            |            |            |      |
|            |            | tatttcgttc |            |            |            |      |
| gcgctgaggt | ctgcctcgtg | aagaaggtgt | tgctgactca | taccaggcct | gaatcgcccc | 5400 |
|            |            | gggagccacg |            |            |            |      |
| gttggtgatt | ttgaactttt | gctttgccac | ggaacggtct | gcgttgtcgg | gaagatgcgt | 5520 |
| gatctgatcc | ttcaactcag | caaaagttcg | atttattcaa | caaagccgcc | gtcccgtcaa | 5580 |
|            |            | gtgttacaac |            |            |            |      |
|            |            | caatttattc |            |            |            |      |
|            |            | aggagaaaac |            |            |            |      |
|            |            | tccgactcgt |            |            |            |      |
|            |            | aagtgagaaa |            |            |            |      |
|            |            | ttctttccag |            |            |            |      |
|            |            | aaccaaaccg |            |            |            |      |
|            |            | aaaaggacaa |            |            |            |      |
|            |            | aacaatattt |            |            |            |      |
|            |            | gatcgcagtg |            |            |            |      |
|            |            | aagaggcata |            |            |            |      |
|            |            |            |            |            |            |      |
| tegggetteg | catcactggc | aacgctacct | agagatast  | ccagaaacaa | ctctggegea | 6300 |
| 209990000  | catacaaccy | atagattgtc | theasett-  | geeegacatt | accgcgagcc | 0360 |
|            |            | agcatccatg |            |            |            |      |
| gtttcccgtt | gaatatggct | cataacaccc | cctgtattac | tgtttatgta | agcagacagt | 6480 |
| tttattgttc | atgatgatat | attttatct  | tgtgcaatgt | aacatcagag | attttgagac | 6540 |
| acaacgtggc | TTCCCCCCC  | ccccattat  | tgaagcattt | atcagggtta | ttgtctcatg | 6600 |
|            |            | tatttagaaa |            |            |            |      |
|            |            | cgtctaagaa | accattatta | tcatgacatt | aacctataaa | 6720 |
| aataggcgta | tcacgaggcc | ctttcgtc   |            |            |            | 6748 |
|            |            |            |            |            |            |      |

<210> 151

```
<211>. 6709
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 151
tegegegtht eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagcttgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tocaacatta cogocatgit gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatqq 420
cocqcctqqc tgaccqccca acgacccccq cccattgacq tcaataatga cqtatqttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttqqcaqtac atctacgtat tagtcatege tattaccatg gtgatgeggt tttqqcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccqcccca ttqacqcaaa tgqqcqgtaq qcqtqtacqq tgqqaqqtct atataaqcaq 900
agetegttta gtgaaccgtc agategcctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacqcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080
tettatgeat getataetgt ttttggettg gggeetatae accceegett cettatgeta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattgqtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteoggtag eggeggaget tecacateeg ageeetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
egegecaeca gacataatag etgacagaet aacagaetgt teettteeat gggtetttte 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acqaccqtqa aqttcacaac qtttqqqcta cccacqcttq cqttccqacc qacccqaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340
ccaacgacat gaacaaagaa atccgtaact gctccttcaa caccaccacc gaaatccgtg 2400
acaaaaaaca gcagggttac gctctgttct accgtccgga catcgttctg ctgaaagaaa 2460
accgtaacaa ctccaacaac tccgaataca tcctgatcaa ctgcaacgct tccaccatca 2520
eccaggettg eccgaaagtt aacttegace egatecegat ceactactge geteeggetg 2580
```

gttacgctat cctgaaatgc aacaacaaaa ccttctccgg taaaggtccg tgcaacaacg 2640 tttccaccgt tcagtgcacc catggtatca aaccggttgt ttccacccag ctgctgctga 2700 acggttccct ggctgaaaaa gaaatcatca tccgttccga aaacctgacc gacaacgtta 2760 aaaccatcat cgttcacctg aacaaatccg ttgaaatcgt ttgcacccgt gctagcgctt 2820 actgcaacat ctccggttcc aaatggaacg aaaccctgaa acgtgttaaa gaaaaactgc 2880 aggaaaacta caacaacaac aaaaccatca aattcgctcc gtcctccggt ggtgacctgg 2940 aaatcaccac ccactccttc aactgeegtg gtgaattett ctactgeaac accgetagec 3000 tgccgtgccg tatcaaacag atcatcaaca tgtggcaggg tgttggtcgt gctatgtacg 3060 ctccgccgat cgctggtaac atcacctgca aatccaacat caccggtctg ctgctggttc 3120 gtgacggtgg tgaagacaac aaaaccgaag aaatcttccg tccgggtggt ggtaacatga 3180 aagacaactg gcgttccgaa ctgtacaaat acaaagttat cgaactgaaa ccgctgggta 3240 tegeteegae eggtgetaag ettacegtte aggetegtea getgetgtee tecategtte 3300 agcagcagtc caacctgctg cgtgctatcg aagctcagca gcacatgctg cagctgaccg 3360 tttggggtat caaacagctg cagacccgtg ttctggctat cgaacgttac ctgaaagacc 3420 agcagetega gatetggaac aacatgacet ggatggaatg ggacegtgaa atetecaact 3480 acaccgacac catctaccgt ctgctggaag actcccagac ccagcaggaa aaaaacgaaa 3540 aagacetget ggetetggae teetggaaaa acetgtggte etggttegae atetecaaet 3600 ggctgtggta catcaaaatc ttcatcatga tcgttggtgg tctgatcggt ctgcgtatca 3660 tettegetgt tetgtecate tgaggateca gatetgetgt geettetagt tgecagecat 3720 ctqttgtttg cccctcccc gtgccttcct tgaccctgga aggtgccact cccactgtcc 3780 tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg 3840 ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc aggcatgctg 3900 gggatgcggt gggctctatg ggtacccagg tgctgaagaa ttgacccggt tcctcctggg 3960 ccagaaagaa gcaggcacat ccccttctct gtgacacacc ctgtccacgc ccctggttct 4020 tagttecage eccaeteata ggacaeteat ageteaggag ggeteegeet teaateceae 4080 ccgctaaagt acttggagcg gtctctccct ccctcatcag cccaccaaac caaacctagc 4140 ctccaagagt gggaagaaat taaagcaaga taggctatta agtgcagagg gagagaaaat 4200 gcctccaaca tgtgaggaag taatgagaga aatcatagaa ttttaaggcc atgatttaag 4260 qccatcatgg ccttaatctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg 4320 gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg 4380 ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa 4440 ggccgcgttg ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg 4500 acgeteaagt cagaggtgge gaaaccegae aggaetataa agataccagg cgtttccccc 4560 tggaagetee etegtgeget etectgttee gaccetgeeg ettaceggat acctgteege 4620 ettteteeet tegggaageg tggegettte teatagetea egetgtaggt ateteagtte 4680 ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg 4740 etgegeetta teeggtaaet ategtettga gteeaaeeeg gtaagaeaeg aettategee 4800 actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga 4860 gttcttgaag tggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc 4920 tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac 4980 caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 5040 atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc 5100 acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa 5160 ttaaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta 5220 ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt 5280 tgcctgactc gggggggggg ggcgctgagg tctgcctcgt gaagaaggtg ttgctgactc 5340 ataccaggee tgaategeee cateatecag ccagaaagtg agggageeac ggttgatgag 5400 agetttgttg taggtggacc agttggtgat tttgaacttt tgctttgcca cggaacggtc 5460 tgcgttgtcg ggaagatgcg tgatctgatc cttcaactca gcaaaagttc gatttattca 5520 acaaagccgc cgtcccgtca agtcagcgta atgctctgcc agtgttacaa ccaattaacc 5580 aattotgatt agaaaaacto atogagoato aaatgaaact goaatttatt catatoagga 5640 ttatcaatac catatttttg aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg 5700 cagttccata ggatggcaag atcctggtat cggtctgcga ttccgactcg tccaacatca 5760 atacaaccta ttaatttccc ctcgtcaaaa ataaggttat caagtgagaa atcaccatga 5820 gtgacgactg aatccggtga gaatggcaaa agcttatgca tttctttcca gacttgttca 5880

PCT/US01/25721

```
WO 02/32943
acaggecage cattacgete gteateaaaa teactegeat caaccaaace gttatteatt 5940
cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt taaaaggaca attacaaaca 6000
ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa 6060
tcaggatatt cttctaatac ctggaatgct gttttcccgg ggatcgcagt ggtgagtaac 6120
catgcatcat caggagtacg gataaaatgc ttgatggtcg gaagaggcat aaattccgtc 6180
agccagttta gtctgaccat ctcatctgta acatcattgg caacgctacc tttgccatgt 6240
ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt cgcacctgat 6300
tgcccgacat tatcgcgagc ccatttatac ccatataaat cagcatccat gttggaattt 6360
aatcgcggcc tcgagcaaga cgtttcccgt tgaatatggc tcataacacc ccttgtatta 6420
ctgtttatgt aagcagacag ttttattgtt catgatgata tatttttatc ttgtgcaatg 6480
taacatcaga gattttgaga cacaacgtgg ctttcccccc ccccccatta ttgaagcatt 6540
tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa 6600
ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 6660
atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtc
<210> 152
<211> 6790
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
```

| gttgcggtgc | tgttaacggt               | ggagggagt  | gtagtctgag                             | cagtactcgt  | tactaccaca  | 1900 |
|------------|--------------------------|------------|--|-------------|-------------|------|
| cgcgccacca | gacataatag               | ctgacagact | aacagactgt                             | tcctttccat  | gggtgttgtg  | 1960 |
| tgcagtcacc | gtcgtcgaca               | catataatca | gatatogo                               | ccactctaga  | gatatogo    | 1000 |
| ccatgcgtgt | tcgtggtatc               | ccacataact | aaccacaata                             | ataastataa  | gatategeea  | 1000 |
| gtttctggat | gatcatcatc               | taccatatta | ttootaacat                             | gtggatgegg  | ggtatttgg   | 7300 |
| ghathccaat | ttggaccgac               | actassaces | cectatteta                             | gegggeeace  | gereactacg  | 2040 |
| accaccatas | anttracear               | gtttgggcta | cccigcccig                             | cgcttcccgac | accaaageet  | 2100 |
| cacaggaaat | agttcacaac<br>cgttctggaa | accettacce | assachtess                             | cgtteegaee  | gacccgaacc  | 2160 |
| ttgaccagat | gcacgaagac               | atcatctcc  | tatacascas                             | cargragaaa  | aacgacatgg  | 2220 |
| aactgacccc | gctgtgcgtt               | accetecece | gergggacea                             | basebbase   | cegtgegtta  | 2280 |
| ccaacgacat | gaacaaagaa               | atcootaact | gcaccaacgc                             | cacciccada  | aacaacgtta  | 2340 |
| acaaaaaaca | gcagggttac               | actetattet | accatacaaa                             | caccaccacc  | gaaatccgtg  | 2400 |
| accotaacaa | ctccaacaac               | tccgataca  | tectastess                             | chacagedet  | togaaagaaa  | 2460 |
| cccagactta | cccgaaagtt               | aacttcgacc | catccat                                | ccactactes  | actocatea   | 2520 |
| gttacgctat | cctgaaatgc               | aacaacaaa  | cetteteeaa                             | taaagataga  | tageness    | 2580 |
| tttccaccat | tcagtgcacc               | catootatca | aaccoattat                             | ttccaccac   | cycaacaacy  | 2640 |
| acggttccct | ggctgaaaaa               | gazatcatca | teesttees                              | anagatanag  | cigetgetga  | 2700 |
| aaaccatcat | cgttcacctg               | aacaaatcca | ttgaaategt                             | thanagast   | gacaacgtta  | 2760 |
| acacccotaa | atccatgcgt               | atcootcoo  | gtgagaggtt                             | chacachage  | ccgaacaaca  | 2820 |
| tcggtgacat | ccatcagage               | tactgcaaca | tetecastta                             | ccaegecaee  | ggtgacatca  | 2880 |
| aacotottaa | ccgtcaggct               | caccycaaca | agaagaagaa                             | caaatggaac  | gaaaccccga  | 2940 |
| catcatcaa  | agaaaaactg<br>tggtgacctg | caggaaaacc | acaacaacaa                             | caaaaccacc  | aaattcgctc  | 3000 |
| tctactccaa | caccgctage               | staccataca | ghatgaaaga                             | caactgeegt  | ggrgaattet  | 3060 |
| atattaatca | tgctatgtac               | actecara   | teactaataa                             | gattateaac  | atgtggcagg  | 3120 |
| tcaccggtct | gctgctggtt               | catalogata | atasaasa                               | Cattactige  | aaacccaaca  | 3180 |
| atccagataa | tggtaacatg               | associates | gegaagacaa                             | caaaaccyaa  | gaaacccccc  | 3240 |
| tcgaactgaa | accgctgggt               | atcoctccoa | coatactaa                              | actycacaaa  | cacaaagtta  | 3300 |
| agctactate | ctccatcgtt               | cancancage | ccaacctcat                             | geetactgee  | caggetegte  | 3360 |
| agcacatgct | gcagctgacc               | atttaggata | tcaacceget                             | gegegetate  | gaagereage  | 3420 |
| tcgaacgtta | cctgaaagac               | caccagggca | agatotogoo                             | geagaceegt  | greetggeta  | 3480 |
| aggaccataa | aatctccaac               | tacaccasca | ccatctacca                             | tataataas   | rggarggaar  | 3540 |
| cccaccaca  | aaaaaacgaa               | aaacacctcc | tagetetage                             | ctcgccggaa  | gacteceaga  | 3600 |
| cctaattcaa | catctccaac               | taactataat | acatcasast                             | cttctggaaa  | aacctgtggt  | 3660 |
| atctaatcaa | tctgcgtatc               | atcttccctc | ttctctccat                             | atanagatag  | accgctggtg  | 3720 |
| taccttctag | ttgccagcca               | tctattattt | acceptees                              | catacattac  | ttasaaataa  | 3/80 |
| aaggtgccac | tcccactgtc               | ctttcctaat | aaaatgagga                             | aattggatgg  | cathatatas  | 3840 |
| ghaggtgtca | ttctattctg               | agaaghaaga | tagaacgagga                            | caccasacca  | cattgtttga  | 3900 |
| aagacaatag | caggcatgct               | 99999-9999 | tagactatat                             | cagcaagggg  | gaggattggg  | 3960 |
| attgacccgg | ttcctcctgg               | accadaaada | aggaggaga                              | teceettete  | grgergaaga  | 4020 |
| cctgtccacg | cccctggttc               | ttagttccag | cccactcat                              | aggagagtga  | tagatacae   | 4080 |
| agactccacc | ttcaatccca               | cccctaaa   | tacttogage                             | aggacaccca  | tageteataa  | 4200 |
| gcccaccaaa | ccaaacctag               | cctccaagag | taaaaaaaa                              | thanageage  | atagggtath  | 4200 |
| aagtgcagag | ggagagaaaa               | tacctccaac | atotoaccaa                             | gtaatgagag  | acayyctact  | 4200 |
| attttaaggc | catgatttaa               | ggccatcatg | accttaatct                             | tecacttest  | cactcactaga | 4320 |
| ctcqctqcqc | teggtegtte               | gactacaaca | agcogtatca                             | actcactcaa  | aggcggtaat  | 4440 |
| acqqttatcc | acagaatcag               | gggataacgc | aggaaagaac                             | atotoaccaa  | aggeggeaat  | 4500 |
| aaaqqccaqq | aaccgtaaaa               | aggccgcgtt | actaacattt                             | thecatage   | teccecce    | 4560 |
| tgacgagcat | cacaaaaatc               | gacgctcaag | tcagaggtgg                             | coasacccoa  | cagactata   | 4500 |
| aagataccag | gcgtttcccc               | ctggaagete | cctcatacac                             | tetectatta  | cascactaca  | 4680 |
| gcttaccqqa | tacctgtccg               | cettteteee | ttcgggaage                             | ataaccettt  | ctcataceta  | 4740 |
| acqctqtaqq | tatctcagtt               | caatataaah | catteactec                             | aagctgggct  | atataceas   | 4800 |
| acccccqtt  | cagcccgacc               | actacaeeth | atccontaac                             | tatcototto  | actocascos  | 40C0 |
| ggtaagacac | gacttatcgc               | cactggcagc | agccachagh                             | aacaggatta  | udeccaacec  | 4000 |
| gtatgtaggc | ggtgctacag               | agttcttgaa | ataataacch                             | aactacggch  | acachagasa  | 4980 |
| aacagtattt | ggtatctgcg               | ctctactaaa | accaattacc                             | ttcggaaaaa  | ganttootac  | 5040 |
| _          |                          |            | J===================================== |             | 2~2~cAAcaA  | JV3U |

```
ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca 5100
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 5160
cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 5220
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 5280
gtaaacttgg tetgacagtt accaatgett aatcagtgag geacetatet cagegatetg 5340
totaltlogt toalcoatag tigocigact oggggggggg gggcgctgag gtctgcctcg 5400
tgaagaaggt gttgctgact cataccaggc ctgaatcgcc ccatcatcca gccagaaagt 5460
gagggagcca cggttgatga gagctttgtt gtaggtggac cagttggtga ttttgaactt 5520
ttgctttgcc acggaacggt ctgcgttgtc gggaagatgc gtgatctgat ccttcaactc 5580
cagtgttaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5700
tgcaatttat tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat 5760
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg 5820
attccgactc qtccaacatc aatacaacct attaatttcc cctcgtcaaa aataaggtta 5880
tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaatggcaa aagcttatgc 5940
atttetttee agaettgtte aacaggecag ceattacget egteateaaa ateaetegea 6000
tcaaccaaac cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg 6060
ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 6120
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg 6180
gggatcgcag tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 6240
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6300
gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 6360
cgatagattg tegeacetga ttgcccgaca ttategegag cecatttata eccatataaa 6420
tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg 6480
ctcataacac cccttgtatt actgtttatg taagcagaca gttttattgt tcatgatgat 6540
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttccccc 6600
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6660
tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct 6720
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6780
                                                                6790
ccctttcgtc
```

```
<211> 6637
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 153
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
```

<210> 153

| WO 02/3294 | 3          |            |            |            | PCT/US                   | 801/257 |
|------------|------------|------------|------------|------------|--------------------------|---------|
| cgtcaatggg | agtttgttt  | ggcaccaaaa | tcaacgggac | tttccaaaat | gtcgtaacaa               | 840     |
|            |            |            |            |            | atataagcag               |         |
|            |            |            |            |            | ttgacctcca               |         |
|            |            |            |            |            | gaacgcggat               |         |
|            |            |            |            |            | ccctttggc                |         |
|            |            |            |            |            | ccttatgcta               |         |
|            |            |            |            |            | accactcccc               |         |
|            |            |            |            |            | caactatctc               |         |
|            |            |            |            |            | tatttttaca               |         |
|            |            |            |            |            | ccccgtgcc                |         |
|            |            |            |            |            | gtgttccgga               |         |
|            |            |            |            |            | ccatgcctcc               |         |
|            |            |            |            |            | acttaggcac               |         |
|            |            |            |            |            | gtatgtgtct               |         |
|            |            |            |            |            | taaggcagcg               |         |
|            |            |            |            |            | ggtaactccc               |         |
|            |            |            |            |            | tgctgccgcg               |         |
|            |            |            |            |            | gggtcttttc               |         |
|            |            |            |            |            | gatatcgcca               |         |
| ccatgcgtgt | tcgtggtatc | ccgcgtaact | ggccgcagtg | gtggatgtgg | ggtatcctgg               | 1980    |
| gtttctggat | gatcatcatc | tgccgtgttg | ttggtaacat | gtgggttacc | gtttactacg               | 2040    |
| gtgttccggt | ttggaccgac | gctaaaacca | ccctgttctg | cgcttccgac | accaaagcct               | 2100    |
| acgaccgtga | agttcacaac | gtttgggcta | cccacgcttg | cgttccgacc | gacccgaacc               | 2160    |
|            |            |            |            |            | aacgacatgg               |         |
| ttgaccagat | gcacgaagac | atcatctccc | tgtgggacca | gtccctgaaa | ccgtgcgtta               | 2280    |
| aactgacccc | gctgtgcgtt | accetgeact | gcaccaacgc | tagccgtaac | tgctccttca               | 2340    |
| acaccaccac | cgaaatccgt | gacaaaaaac | agcagggtta | cgctctgttc | taccgtccgg               | 2400    |
|            |            |            |            |            | atcctgatca               |         |
|            |            |            |            |            | ccgatcccga               |         |
|            |            |            |            |            | accttctccg               |         |
|            |            |            |            |            | aaaccggttg               |         |
|            |            |            |            |            | atccgttccg               |         |
|            |            |            |            |            | gttgaaatcg               |         |
|            |            |            |            |            | gaaaccctga               |         |
|            | · ·        |            |            |            | aaattcgctc               |         |
|            |            |            |            |            | ggtgaattct               |         |
|            |            |            |            |            | accatcaccc               |         |
|            |            |            |            |            | gctatgtacg               |         |
|            |            |            |            |            | ctgctggttc               |         |
|            |            |            |            |            | ggtaacatga               |         |
|            |            |            |            |            | ccgctgggta               |         |
|            |            |            |            |            | tccatcgttc               |         |
| ageageagee | caaccigcig | egigerateg | aageteagea | geacatgetg | cagctgaccg               | 3360    |
|            |            |            |            |            | ctgaaagacc               |         |
|            |            |            |            |            | atctccaact               |         |
|            |            |            |            |            | aaaaacgaaa               |         |
|            |            |            |            |            | atctccaact               |         |
|            |            |            |            |            | gttgtttgcc<br>tcctaataaa |         |
|            |            |            |            |            | ggtggggtgg               |         |
|            |            |            |            |            | gatgcggtgg               |         |
|            |            |            |            |            | agaaagaagc               |         |
|            |            |            |            |            | gttccagccc               |         |
|            |            |            |            |            | gctaaagtac               |         |
|            |            |            |            |            | ccaagagtgg               |         |
| 5-550      |            |            |            |            |                          |         |

```
gaagaaatta aagcaagata ggctattaag tgcagaggga gagaaaatgc ctccaacatg 4140
tgaggaagta atgagagaaa tcatagaatt ttaaggccat gatttaaggc catcatggcc 4200
ttaatettee getteetege teactgaete getgegeteg gtegttegge tgeggegage 4260
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4320
aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 4380
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4440
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4500
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 4560
gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt 4620
tegetecaag etgggetgtg tgeacgaace eccegtteag eccgaeeget gegeettate 4680
eggtaactat egtettgagt ecaaceeggt aagacaegae ttategeeae tggcageage 4740
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 4800
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 4860
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4920
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaaqaaqa 4980
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 5040
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5100
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 5160
cagtgaggca cotatotoag ogatotgtot atttogttoa tocatagttg cotgactogg 5220
ggggggggg cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat accaggcctg 5280
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttgttgta 5340
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcggg 5400
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5460
tecegteaag teagegtaat getetgeeag tgttacaace aattaaceaa ttetgattag 5520
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5580
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5640
atggcaagat cetggtateg gtetgegatt cegactegte caacateaat acaacetatt 5700
aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5760
teeggtgaga atggeaaaag ettatgeatt tettteeaga ettgtteaac aggeeageea 5820
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 5880
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 5940
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 6000
tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tqcatcatca 6060
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccqtcag ccagtttagt 6120
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6180
tetggegeat egggetteee atacaatega tagattgteg cacetgattg eeegacatta 6240
tegegagece atttatacee atataaatea geateeatgt tggaatttaa tegeggeete 6300
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6360
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6420
ttttgagaca caacgtggct ttcccccccc ccccattatt gaagcattta tcagggttat 6480
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6540
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6600
acctataaaa ataggcgtat cacgaggccc tttcgtc
```

```
<210> 154
<211> 6670
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 154
```

| tegegegttt | cggtgatgac | ggtgaaaacc | tctgacacat | gcagctcccg | gagacggtca | 60 - |
|------------|------------|------------|------------|------------|------------|------|
| cagcttgtct | gtaagcggat | gccgggagca | gacaagcccg | tcagggcgcg | tcagcgggtg | 120  |
| ttggcgggtg | teggggetgg | cttaactatg | cggcatcaga | gcagattgta | ctgagagtgc | 180  |
| accatatgcg | gtgtgaaata | ccgcacagat | gcgtaaggag | aaaataccgc | atcagattgg | 240  |
| ctattggcca | ttgcatacgt | tgtatccata | tcataatatg | tacatttata | ttggctcatg | 300  |
| tccaacatta | ccgccatgtt | gacattgatt | attgactagt | tattaatagt | aatcaattac | 360  |
| ggggtcatta | gttcatagcc | catatatgga | gttccgcgtt | acataactta | cggtaaatgg | 420  |
|            |            |            |            |            | cgtatgttcc |      |
|            |            |            |            |            | tacggtaaac |      |
|            |            |            |            |            | ttgacgtcaa |      |
| tgacggtaaa | tggcccgcct | ggcattatgc | ccagtacatg | accttatggg | actttcctac | 660  |
|            |            |            |            |            | tttggcagta |      |
| catcaatggg | cgtggatagc | ggtttgactc | acggggattt | ccaagtctcc | accccattga | 780  |
| cgtcaatggg | agtttgtttt | ggcaccaaaa | tcaacgggac | tttccaaaat | gtcgtaacaa | 840  |
| ctccgcccca | ttgacgcaaa | tgggcggtag | gcgtgtacgg | tgggaggtct | atataagcag | 900  |
|            |            |            |            |            | ttgacctcca |      |
| tagaagacac | cgggaccgat | ccagcctccg | cggccgggaa | cggtgcattg | gaacgcggat | 1020 |
| tccccgtgcc | aagagtgacg | taagtaccgc | ctatagactc | tataggcaca | cccctttggc | 1080 |
| tcttatgcat | gctatactgt | ttttggcttg | gggcctatac | acccccgctt | ccttatgcta | 1140 |
| taggtgatgg | tatagcttag | cctataggtg | tgggttattg | accattattg | accactcccc | 1200 |
| tattggtgac | gatactttcc | attactaatc | cataacatgg | ctctttgcca | caactatctc | 1260 |
| tattggctat | atgccaatac | tctgtccttc | agagactgac | acggactctg | tatttttaca | 1320 |
| ggatggggtc | ccatttatta | tttacaaatt | cacatataca | acaacgccgt | ccccgtgcc  | 1380 |
| cgcagttttt | attaaacata | gcgtgggatc | tccacgcgaa | tetegggtae | gtgttccgga | 1440 |
| catgggctct | tctccggtag | cggcggagct | tccacatccg | agccctggtc | ccatgcctcc | 1500 |
| agcggctcat | ggtcgctcgg | cagctccttg | ctcctaacag | tggaggccag | acttaggcac | 1560 |
| agcacaatgc | ccaccaccac | cagtgtgccg | cacaaggccg | tggcggtagg | gtatgtgtct | 1620 |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact | taaggcagcg | 1680 |
|            |            |            |            |            | ggtaactccc |      |
|            |            |            |            |            | tgctgccgcg |      |
|            |            |            |            |            | gggtcttttc |      |
| tgcagtcacc | gtcgtcgaca | cgtgtgatca | gatatcgcgg | ccgctctaga | gatatcgcca | 1920 |
| ccatgcgtgt | tcgtggtatc | ccgcgtaact | ggccgcagtg | gtggatgtgg | ggtatcctgg | 1980 |
| gtttctggat | gatcatcatc | tgccgtgttg | ttggtaacat | gtgggttacc | gtttactacg | 2040 |
|            |            |            |            |            | accaaagcct |      |
|            |            |            |            |            | gacccgaacc |      |
|            |            |            |            |            | aacgacatgg |      |
|            |            |            |            |            | ccgtgcgtta |      |
|            |            |            |            |            | tgctccttca |      |
|            |            |            |            |            | taccgtccgg |      |
| acatcgttct | gctgaaagaa | aaccgtaaca | actccaacaa | ctccgaatac | atcctgatca | 2460 |
| actgcaacgc | ttccaccatc | acccaggett | gcccgaaagt | taacttcgac | ccgatcccga | 2520 |
|            |            |            |            |            | accttctccg |      |
|            |            |            |            |            | aaaccggttg |      |
|            |            |            |            |            | atccgttccg |      |
| aaaacctgac | cgacaacgtt | aaaaccatca | tcgttcacct | gaacaaatcc | gttgaaatcg | 2760 |
|            |            |            |            |            | gaaaccctga |      |
|            |            |            |            |            | aaattcgctc |      |
|            |            |            |            |            | ggtgaattct |      |
|            |            |            |            |            | atgtggcagg |      |
|            |            |            |            |            | aaatccaaca |      |
| tcaccggtct | gctgctggtt | cgtgacggtq | gtgaagacaa | caaaaccgaa | gaaatcttcc | 3120 |
|            |            |            |            |            | tacaaagtta |      |
|            |            |            |            |            | caggetegte |      |
|            |            |            |            |            | gaageteage |      |
|            | -          |            | <b>.</b>   |            |            |      |

| agcacatgct | gcagctgacc | gtttggggta | tcaaacagct | gcagacccgt | gttctggcta  | 3360 |
|------------|------------|------------|------------|------------|-------------|------|
| tcgaacgtta | cctgaaagac | cagcagctcg | agatctggaa | caacatgacc | tggatggaat  | 3420 |
| gggaccgtga | aatctccaac | tacaccgaca | ccatctaccg | tctgctggaa | gactcccaga  | 3480 |
| cccagcagga | aaaaaacgaa | aaagacctgc | tggctctgga | ctcctggaaa | aacctgtggt  | 3540 |
| cctggttcga | catctccaac | tggctgtggt | acatcaaaat | cttcatcatq | atcottooto  | 3600 |
| gtctgatcgg | tctgcgtatc | atcttcgctg | ttctgtccat | ctgaggatcc | agatetgetg  | 3660 |
| tgccttctag | ttgccagcca | tetattatt  | gcccctcccc | catacettee | ttgaccctgg  | 3720 |
| aaggtgccac | tcccactgtc | ctttcctaat | aaaatgagga | aattgcatcg | cattototoga | 3780 |
| gtaggtgtca | ttctattctg | gagaataaga | tggggcagga | cadcaaddd  | gaggattggg  | 3840 |
| aagacaatag | caggcatgct | ggggatgcgg | tagactctat | gggtaccag  | atactassas  | 3900 |
| attgacccgg | ttcctcctgg | qccaqaaaqa | agcaggcaca | tecettete  | tataacacac  | 3960 |
| cctqtccacq | cccctggttc | ttagttccag | cccactcat  | aggaractra | tageteagga  | 4020 |
| qqqctccqcc | ttcaatccca | cccactaaaa | tacttogage | agtetetee  | teceteatea  | 4080 |
| qcccaccaaa | ccaaacctag | cctccaagag | taggaagaaa | ttaaagcaag | ataggetatt  | 4140 |
| aagtgcagag | ggagagaaaa | tocctccaac | atotoaggaa | gtaatgagag | asatcataca  | 4200 |
| attttaaggc | catgatttaa | ggccatcatg | gccttaatct | tecacttect | cactcactga  | 4260 |
| ctcactacac | teggtegtte | gactacaaca | agcoctatce | actcactcaa | agggggtaat  | 4220 |
| acoottatco | acagaatcag | addataacac | ageggeacea | atataaaaa  | aggeggeaat  | 4320 |
| aaaaaccaaa | aaccotaaaa | aggacaacgc | actacattt  | ttacatacaa | haagaaaaa   | 4380 |
| taacaaacat | aaccgtaaaa | aggeegegee | tananata   | gannaga    | Leegececee  | 4440 |
| agacyagcac | cacaaaaatc | stanceta   | ccagaggtgg | cgaaacccga | caggactata  | 4500 |
| acttaccas  | gcgtttcccc | ccggaagece | ttaaaaaaa  | teteetgtte | cgaccetgee  | 4550 |
| accetatora | tacctgtccg | ccccccccc  | ccegggaage | grggegerre | ctcatagete  | 4620 |
| acgeegeagg | tatctcagtt | cggrgraggr | cyccegetee | aagetggget | grgrgcacga  | 4680 |
| acteccegee | cagcccgacc | gergegeerr | acceggtaac | categicitg | agtccaaccc  | 4740 |
| ggcaagacac | gacttatcgc | cactggcage | ageeaetggt | aacaggatta | gcagagcgag  | 4800 |
| grargragge | ggtgctacag | agetetegaa | grggrggeer | aactacggct | acactagaag  | 4860 |
| ababbasta  | ggtatctgcg | ccccgccgaa | gecagetace | ttcggaaaaa | gagttggtag  | 4920 |
| cittigatee | ggcaaacaaa | ccaeegetgg | cageggegge | ttttttgttt | gcaagcagca  | 4980 |
| gartacgege | agaaaaaaag | gateteaaga | agateette  | accetecta  | cggggtctga  | 5040 |
| cgctcagtgg | aacgaaaact | cacgitaagg | gattttggtt | atgagattat | caaaaaggat  | 5100 |
| cttcacctag | atccttttaa | accaaaaacg | aagttttaaa | tcaatctaaa | gtatatatga  | 5160 |
| gcaaacttgg | tctgacagtt | accaatgett | aatcagtgag | gcacctatct | cagcgatctg  | 5220 |
| Lecatteegt | tcatccatag | ttgcctgact | caaaaaaaaa | gggcgctgag | gtetgeeteg  | 5280 |
| tgaagaaggt | gttgctgact | cataccagge | ctgaatcgcc | ccatcatcca | gccagaaagt  | 5340 |
| gagggagcca | cggttgatga | gagetttgtt | gtaggtggac | cagttggtga | ttttgaactt  | 5400 |
| ttgctttgcc | acggaacggt | ctgcgttgtc | gggaagatgc | gtgatctgat | ccttcaactc  | 5460 |
| agcaaaagtt | cgatttattc | aacaaagccg | ccgtcccgtc | aagtcagcgt | aatgctctgc  | 5520 |
| cagtgttaca | accaattaac | caattctgat | tagaaaaact | catcgagcat | caaatgaaac  | 5580 |
| tgcaatttat | tcatatcagg | attatcaata | ccatatttt  | gaaaaagccg | tttctgtaat  | 5640 |
| gaaggagaaa | actcaccgag | gcagttccat | aggatggcaa | gatcctggta | teggtetgeg  | 5700 |
| attccgactc | gtccaacatc | aatacaacct | attaatttcc | cctcgtcaaa | aataaggtta  | 5760 |
| tcaagtgaga | aatcaccatg | agtgacgact | gaatccggtg | agaatggcaa | aagcttatgc  | 5820 |
| atttetttee | agacttgttc | aacaggccag | ccattacgct | cgtcatcaaa | atcactcgca  | 5880 |
| tcaaccaaac | cgttattcat | tcgtgattgc | gcctgagcga | gacgaaatac | gcgatcgctg  | 5940 |
| ttaaaaggac | aattacaaac | aggaatcgaa | tgcaaccggc | gcaggaacac | tgccagcgca  | 6000 |
| tcaacaatat | tttcacctga | atcaggatat | tcttctaata | cctggaatgc | tgttttcccg  | 6060 |
| gggatcgcag | tggtgagtaa | ccatgcatca | tcaggagtac | ggataaaatg | cttgatggtc  | 6120 |
| ggaagaggca | taaattccgt | cagccagttt | agtctgacca | tctcatctgt | aacatcattg  | 6180 |
| gcaacgctac | ctttgccatg | tttcagaaac | aactctggcg | catcgggctt | cccatacaat  | 6240 |
| cgatagattg | tegeacetga | ttgcccgaca | ttatcgcgag | cccatttata | cccatataaa  | 6300 |
| tcagcatcca | tgttggaatt | taatcgcggc | ctcgagcaag | acgtttcccg | ttgaatatgg  | 6360 |
| ctcataacac | cccttgtatt | actgtttatg | taagcagaca | gttttattgt | tcatgatgat  | 6420 |
| atattttat  | cttgtgcaat | gtaacatcag | agattttgag | acacaacgtg | gctttccccc  | 6480 |
| ccccccatt  | attgaagcat | ttatcagggt | tattgtctca | tgagcggata | catatttgaa  | 6540 |
| tgtatttaga | aaaataaaca | aataggggtt | ccgcgcacat | ttccccgaaa | agtgccacct  | 6600 |
|            |            |            |            |            |             |      |

gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6660 ccctttcgtc 6670

<210> 155

```
<211> 6547
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 155
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattqa 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eccetttgge 1080
tottatgcat gotatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500
ageggeteat ggtegetegg eageteettg etectaacag tggaggecag acttaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc getgtgcgtt accetgcact gcaccaacgc tagcatcaac tgcaacgctt 2340
ccaccatcac ccaggettge ccgaaagtta acttegacee gatecegate cactactgeg 2400
ctccggctgg ttacgctatc ctgaaatgca acaacaaaac cttctccggt aaaggtccgt 2460
```

| ggaagaage    | ++         |            |                          |            | L          |      |
|--------------|------------|------------|--------------------------|------------|------------|------|
|              |            |            | atggtatcaa               |            |            |      |
|              |            |            | aaatcatcat               |            |            |      |
|              |            |            | acaaatccgt               |            |            |      |
|              |            |            | tcggtccggg               |            |            |      |
|              |            |            | actgcaacat               |            |            |      |
|              |            |            | aggaaaacta               |            |            |      |
|              |            |            | aaatcaccac               |            |            |      |
|              |            |            | tgccgtgccg               |            |            |      |
|              |            |            | ctccgccgat               |            |            |      |
| aatccaacat   | caccggtctg | ctgctggttc | gtgacggtgg               | tgaagacaac | aaaaccgaag | 3060 |
| aaatcttccg   | tccgggtggt | ggtaacatga | aagacaactg               | gcgttccgaa | ctgtacaaat | 3120 |
| acaaagttat   | cgaactgaaa | ccgctgggta | tegeteegae               | cggtgctaag | cttaccgttc | 3180 |
|              |            |            | agcagcagtc               |            |            |      |
|              |            |            | tttggggtat               |            |            |      |
|              |            |            | agcagctcga               |            |            |      |
|              |            |            | acaccgacac               |            |            |      |
|              |            |            | aagacctgct               |            |            |      |
|              |            |            | ggctgtggtg               |            |            |      |
|              |            |            | cctccccgt                |            |            |      |
| gtgccactcc   | cactatectt | tectaataaa | atgaggaaat               | tacatcacat | tatataaata | 3660 |
|              |            |            | ggcaggacag               |            |            |      |
|              |            |            | gctctatggg               |            |            |      |
|              |            |            | aggcacatcc               |            |            |      |
|              |            |            | cactcatagg               |            |            |      |
|              |            |            |                          |            |            |      |
|              |            |            | ttggagcggt               |            |            |      |
|              |            |            | gaagaaatta               |            |            |      |
|              |            |            | tgaggaagta               |            |            |      |
| actacactac   | gacctaaggc | taccacggcc | ttaatcttcc               | getteetege | cactgaete  | 4140 |
| gttatccaca   | geegeeegge | ataaccaaca | ggtatcagct               | tasaasaasa | cygtaatacg | 4200 |
| gccacccaca   | catasasaa  | ccacattact | aaagaacatg<br>ggcgtttttc | catacastas | gccagcaaaa | 4200 |
| caaccatcac   | aaaaatccac | actosactos | gaggtggcga               | aacaggetee | geeceectga | 4320 |
| ataccaccac   | tttcccctc  | geteaagtea | cgtgcgctct               | aaccegacag | gactacaaag | 4380 |
|              |            |            |                          |            |            |      |
| ctatagatat   | ctcagttgcc | tatacateat | gggaagcgtg               | gegetttete | tagereacy  | 4500 |
|              |            |            | tcgctccaag               |            |            |      |
|              |            |            | cggtaactat               |            |            |      |
| tatagacacgac | gatagagaga | tggcagcagc | cactggtaac               | aggactagca | gagcgaggta | 4680 |
|              |            |            | gtggcctaac               |            |            |      |
|              |            |            | agttaccttc               |            |            |      |
| ttgatcegge   | aaacaaacca | cegetggtag | cggtggtttt               | tttgtttgca | agcagcagat | 4860 |
| Lacgegeaga   | aaaaaaggat | ctcaagaaga | tcctttgatc               | ttttctacgg | ggtctgacgc | 4920 |
| tcagtggaac   | gaaaactcac | gttaagggat | tttggtcatg               | agattatcaa | aaaggatctt | 4980 |
| cacctagate   | cttttaaatt | aaaaatgaag | ttttaaatca               | atctaaagta | tatatgagta | 5040 |
| aacttggtct   | gacagttacc | aatgcttaat | cagtgaggca               | cctatctcag | cgatctgtct | 5100 |
| atttegttea   | tccatagttg | cctgactcgg | aaaaaaaaaa               | cgctgaggtc | tgcctcgtga | 5160 |
|              |            |            | aatcgcccca               |            |            |      |
| ggagccacgg   | ttgatgagag | ctttgttgta | ggtggaccag               | ttggtgattt | tgaacttttg | 5280 |
| ctttgccacg   | gaacggtctg | cgttgtcggg | aagatgcgtg               | atctgatcct | tcaactcagc | 5340 |
| aaaagttcga   | tttattcaac | aaagccgccg | tcccgtcaag               | tcagcgtaat | gctctgccag | 5400 |
| tgttacaacc   | aattaaccaa | ttctgattag | aaaaactcat               | cgagcatcaa | atgaaactgc | 5460 |
|              |            |            | tatttttgaa               |            |            |      |
|              |            |            | atggcaagat               |            |            |      |
|              |            |            | aatttcccct               |            |            |      |
|              |            |            | tccggtgaga               |            |            |      |
| tctttccaga   | cttgttcaac | aggccagcca | ttacgctcgt               | catcaaaatc | actcgcatca | 5760 |

```
accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 5820
aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 5880
acaatatttt cacctgaatc aggatatct tctaatacct ggaatgctgt tttcccgggg 5940
atcgcagtgg tgagtaacca tgcatcatca ggagtacgga taaaaatgctt gatggtcgga 6000
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6060
acgctacctt tgccatgtt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6120
tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atatagacta 6180
gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg aatatggctc 6240
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgtca tgatgatata 6300
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct tccccccc 6360
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6420
atttagaaaa ataaacaaat aggggttccg cgcacattc cccgaaaagt caccgaggccc 6540
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6540
tttcgtc
```

<210> 156 <211> 6679 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgacg taagtacege ctatagacte tataggeaca eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac accecegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

|            |            |             |            |            | tgctgccgcg |      |
|------------|------------|-------------|------------|------------|------------|------|
|            |            |             |            |            | gggtcttttc |      |
|            |            |             |            |            | gatategeca |      |
|            |            |             |            |            | ggtatcctgg |      |
|            |            |             |            |            | gtttactacg |      |
|            |            |             |            |            | accaaagcct |      |
|            |            |             |            |            | gacccgaacc |      |
|            |            |             |            |            | aacgacatgg |      |
|            |            |             |            |            | ccgtgcgtta |      |
|            |            |             |            |            | tgctccttca |      |
| acaccaccac | cgaaatccgt | gacaaaaaac  | agcagggtta | cgctctgttc | taccgtccgg | 2400 |
|            |            |             |            |            | atcctgatca |      |
| actgcaacgc | ttccaccatc | acccaggctt  | gcccgaaagt | taacttcgac | ccgatcccga | 2520 |
| tccactactg | cgctccggct | ggttacgcta  | tcctgaaatg | caacaacaaa | accttctccg | 2580 |
|            |            |             |            |            | aaaccggttg |      |
|            |            |             |            |            | atccgttccg |      |
|            |            |             |            |            | gttgaaatcg |      |
|            |            |             |            |            | ggtcagacct |      |
| tctacgctac | cggtgacatc | atcggtgaca  | tccgtcaggc | ttactgcaac | atctccggtt | 2880 |
| ccaaatggaa | cgaaaccctg | aaacgtgtta  | aagaaaaact | gcaggaaaac | tacaacaaca | 2940 |
| acaaaaccat | caaattcgct | ccgtcctccg  | gtggtgacct | ggaaatcacc | acccactcct | 3000 |
| tcaactgccg | tggtgaattc | ttctactgca  | acaccgctag | cctgccgtgc | cgtatcaaac | 3060 |
| agatcatcaa | catgtggcag | ggtgttggtc  | gtgctatgta | cgctccgccg | atcgctggta | 3120 |
| acatcacctg | caaatccaac | atcaccggtc  | tgctgctggt | tcgtgacggt | ggtgaagaca | 3180 |
|            |            |             |            |            | tggcgttccg |      |
|            |            |             |            |            | accggtgcta |      |
|            |            |             |            |            | tccaacctgc |      |
|            |            |             |            |            | atcaaacagc |      |
|            |            |             |            |            | gagatctgga |      |
|            |            |             |            |            | accatctacc |      |
|            |            |             |            |            | ctggctctgg |      |
|            |            |             |            |            | tgaggatcca |      |
|            |            |             |            |            | gtgccttcct |      |
|            |            |             |            |            | attgcatcgc |      |
|            |            |             |            |            | agcaaggggg |      |
|            |            |             |            |            | ggtacccagg |      |
|            |            |             |            |            | ccccttctct |      |
|            |            |             |            |            | ggacactcat |      |
|            |            |             |            |            | gtctctccct |      |
|            |            |             |            |            | taaagcaaga |      |
|            |            |             |            |            | taatgagaga |      |
|            |            |             |            |            | ccgcttcctc |      |
| gctcactgac | tegetgeget | cggtcgttcg  | gctgcggcga | gcggtatcag | ctcactcaaa | 4320 |
|            |            |             |            |            | tgtgagcaaa |      |
|            |            |             |            |            | tccataggct |      |
|            |            |             |            |            | gaaacccgac |      |
|            |            |             |            |            | ctcctgttcc |      |
|            |            |             |            |            | tggcgctttc |      |
|            |            |             |            |            | agctgggctg |      |
|            |            |             |            |            | atcgtcttga |      |
|            |            |             |            |            | acaggattag |      |
|            |            |             |            |            | actacggcta |      |
|            |            |             |            |            | tcggaaaaag |      |
|            |            |             |            |            | tttttgtttg |      |
| Caadcadcad | attacococa | gaaaaaaaagg | atctcaacaa | gatectttea | tetttetae  | 5040 |
| Jurguaguag |            | 553         |            | ,          |            |      |

```
ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc 5100
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag 5160
tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 5220
tetqueteqt gaagaaggtg ttgetgacte ataccaggee tgaategeee cateatecag 5340
ccaqaaaqtq aqqqaqccac qqttqatqaq aqctttqttq taggtggacc agttggtgat 5400
cttcaactca qcaaaaqttc qatttattca acaaaqccqc cqtcccqtca agtcagcqta 5520
atgetetgee agtgttacaa ecaattaace aattetgatt agaaaaacte ategageate 5580
aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5640
ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag atcctggtat 5700
cqqtctqcqa ttccgactcg tccaacatca atacaaccta ttaatttccc ctcgtcaaaa 5760
ataaqqttat caaqtqagaa atcaccatga gtgacqactg aatccggtga gaatggcaaa 5820
aqcttatqca tttctttcca gacttgttca acaggccagc cattacgctc gtcatcaaaa 5880
tractograt caaccaaacc gttattcatt cgtgattgcg cctgagcgag acgaaatacg 5940
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact 6000
gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6060
gttttcccgg ggatcgcagt ggtgagtaac catgcatcat caggagtacg gataaaatgc 6120
ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat ctcatctgta 6180
acatcattqq caacqctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc 6240
ccatacaatc qataqattgt cgcacctgat tgcccgacat tatcgcgagc ccatttatac 6300
ccatataaat caqcatccat qttqqaattt aatcqcqqcc tcqaqcaaqa cqtttcccqt 6360
tqaatatqqc tcataacacc ccttgtatta ctgtttatgt aagcagacag ttttattgtt 6420
catqatqata tatttttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6480
ctttccccc cccccatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 6540
atatttqaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 6600
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6660
                                                             6679
atcacgagge cetttegte
```

```
<210> 157
<211> 6438
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegtht eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttqqcqqqtq tcqqqqctqq cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
qqqqtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccqcctggc tqaccqccca acqacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg qcaqtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacqqtaaa tqqcccqcct qqcattatqc ccagtacatg accttatggg actttcctac 660
ttqqcagtac atctacqtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acgggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccqcccca ttqacqcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
```

| agetegttta gtgaacegte agategeetg gagacgeeat ceaegetgtt ttgaceteea 960   |             |
|---|-------------|
| agetegttta gtgaacegte agategeetg gagaegeete total agaegeggat 1020 tagaagaeae egggaeegat ceageeteeg eggeegggaa eggtgeattg gaacgeggat 1020 tagaagaeae egggaeegat eestetgge 1080   |             |
| tagaagacac cgggaccgat ccagcctccg cggcatc tataggcaca ccctttggc 1080  |             |
| tagaagacac cgggaccgat ccagectecg cggccsssad cssss cccctttggc 1080 teecegtgec aagagtgacg taagtaccgc ctatagacte tataggcaca cccctttggc 1080  |             |
| teccegtgee aagagtgacg taagtacege teatagates accecegett cettatgeta 1140 tettatgeat getataetgt tettagettg gggeetatae accecegett cettatgeta 1200   |             |
| tettatgeat getatactgt tittggettg gggtetattd accattattg accactecce 1200 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactatete 1260  | J           |
| taggtgatgg tatagettag ectataggtg tgggetateg actatete 1260 tattggtgac gatactttec attactaate cataacatgg etetttgeca caactatete 1260 tattggtgac gatactttec attactaate cataacatgg etetttgeca tattttaca 1320  | J           |
|   |             |
| tattggctat atgccaatac tetgteette agagacegat tattggctat atgccaatac tetgteette agagacegat tattggctat atgccaatac tetgteette agagacegat tattggctat atgcccgtgcc 1380 ggattggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 ggattggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 | )           |
| ggatggggtc ccatttatta tttacaaatt tatattata totogggtac gtgttccgga 1440 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440  | )           |
| cgcagttttt attaaacata gcgtgggatt tccacatccg agccctggtc ccatgcctcc 1500 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500   | )           |
| catgggetet teteeggtag eggeggaget teeatateeg tygaggeeag acttaggeac 1560 ageggeteat ggtegetegg eageteettg eteetaacag tygaggeeag acttaggeac 1620   | )           |
| ageggeteat ggtegetegg cageteettg eteetaatug eggtgggtagg gtatgtgtet 1620 ageacaatge ecaceaccae cagtgtgeeg cacaaggeeg tggeggtagg gtatgtgtet 1680  | )           |
| agcacaatgc ccaccaccac cagtgtgccg cacaaggccg cssossoss sagcagaged 1680 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagatcaga ggtaactccc 1740  | 3           |
| gaaaatgagc gtggagattg ggctcgcacg gctgatgcag acggatcaga ggtaactccc 1740 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1800   | 0           |
| gcagaagaag atgcaggcag ctgagttgtt gtattetgat tagtstegt tgctgccgcg 1800 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  | 0           |
| gttgcggtgc tgttaacggt ggagggcagt gtagtctgts object gggtcttttc 1860 cgcgccacca gacataatag ctgacagact aacagactgt tectttccat gggtcttttc 1860 cgcgccacca gacataatag ctgacagact gggagctgt gtgagcggcg gcgagctgga 1920   | 0           |
| cgcgccacca gacataatag ctgacagact aacagacess totageggcg gcgagetgga 1926<br>tgcagtcacc gtcgtcgaca tgggcgcccg cgccagcgtg ctgageggcg gcgagetgga 1926  | 0           |
| tgcagtcacc gtcgtcgaca tgggcgcccg cgccagaag aagtacaagc tgaagcacat 198 ccgctgggag aagatccgcc tgcgcccgg cggcaagaag aagtacaagc tgaagcacag 204   | 0           |
| ccgctgggag aagatccgcc tgcgccccgg tggcaagaag ccggcctgc tggagaccag 204 cgtgtgggcc agccgcgagc tggagcgctt cgccgtgaac cccggcctgc tggagaccag 210  | 0           |
| cgtgtgggcc agccgcgagc tggagcgctt tggeggada tropped cgagcgagcagc tggagccagc ctgcagaccg gcagcgagga 210 cgagggctgc cgccagatcc tgggccagct gcagccagc gcagcgagcagc gcatcgagat 216   | 0           |
| cgagggctgc cgccagatec tgggccagct gcagtccagc ccgcaccagc gcatcgagat 216 gctgcgcagc ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgagat 222   | 0           |
| getgegeage etgtacaaca eegtggeead eetgtacege gegaacaaga geaagaagaa 222 caaggacacc aaggaggee tggacaagat egaggaggag cagaacaaga geaagaagaa 228  | 0           |
| caaggacacc aaggaggeec tggacaagat egaggaggag eagannan actaccccat 228 ggeccagcag geegeegeeg acaceggeea cagcaaccag gtgagccaga actaccccat 228   | 0.          |
| ggcccagcag gccgccgccg acaccggcca cagcaccag gggcccccgca ccctgaacgc 234 cgtgcagaac atccagggcc agatggtgca ccaggccatc agccccgag gtgatcccca tgttcagcgc 240   | 0           |
| cgtgcagaac atccagggcc agatggtgca ctaggccatb agatcccca tgttcagcgc 240 ctgggtgaag gtggtggagg agaaggcctt cagccccgag gtgatcccca tgttcagcgc 240  | 0           |
| ctgggtgaag gtggtggagg agaaggeett cageeeegag gtgacaccg tgggeggeea 246 cetgagegag ggegeeacce eccaggacet gaacaccatg etgaacaccg tgggeggeea 252  | 20          |
| cctgagcgag ggcgccaccc cccaggacct gaacaccagg object agtgggaccg 252 ccaggccgc atgcagatgc tgaaggagac catcaacgag gaggccgcag agtgggaccg 252  | 30          |
| ccaggeegee atgeagatge tgaaggagae tattaatgug guggegage cceggggag 258 cgtgcaccce gtgcacgeeg geeccatege ecceggecag atgegegage ccageggeag 268   | 10          |
| cgtgcacccc gtgcaccacca gcaccctgca ggagcagatc ggctggatga ccaacaaccc 264 cgacatcgcc ggcaccacca gcaccctgca ggagcagatc ctggatcatc ctgggcctga acaagatcgt 270   | 00          |
|   |             |
| ccccatccc gtgggcgaga tctacaagcg ctggatcado cogga agcccttccg 276 gcgcatgtac agccccacca gcatcctgga catccgccag ggccccaagg agcccttccg 286   | 20          |
| gcgcatgtac agecccacca gcatcetgga catteggag gagccagcc aggaggtgaa 282 cgactacgtg gaccgcttct acaagaccct gcgcgccgag caggccagcc aggaggtgaa 282   | 80          |
| cgactacgtg gaccgcttct acaagaccct gcgcgccgag caggactgca agaccatcct 280 gaactggatg accgagaccc tgctggtgca gaacgccaac cccgactgca agaccatcct 280 gaactggatg accgcctgcc agggcgtggg 290  | 40          |
| gaactggatg accgagaccc tgctggtgda gaacgctddd ddegardy<br>gaaggccctg ggccccgccg ccaccctgga ggagatgatg accgcctgcc agggcgtggg 29-   | 00          |
| gaaggeeetg ggeeegeeg ceaceetgga ggagaegaeg ageeaggtga ceaacagege 30 eggeeeegge cacaaggeee gegtgetgge egaggeeatg ageaggtga agtgetteaa 30   | 60          |
| caccatcatg atgcagegeg geaactteeg caaccagege aagategtga agtgctteaa 30 caccatcatg atgcagegeg geaactteeg caccagege ecceqeaaga agggetgetg 31  | 20          |
| caccatcatg atgragegeg geaactiveg caactagege days aggreen 31 ctgeggeaag gagggeeaca eegeegeaa etgeegege eeeegeaaga agggetgetg 31  | ВО          |
| ctgcggcaag gagggccaca ccgcccgcaa ctgccgcgc btts. caggggacagg ctaattttt 31 gaagtgcggc aaggagggc accagatgaa ggactgcacc gagcgacagg ctaattttt 31  | 40          |
| agggaagate tggcettece acaagggaag gebbbaggaa gagacaacaa etecetetea 33  | 00          |
| gccaacagcc ccaccagaag agagetteag tectosagat tecttagat cactettigg 33   | 60          |
| gaagcaggag ccgatagaca aggaactgta tcctttagct tccagaccat tctagaccag 34  | 20          |
| gaagcaggag ccgatagaca aggaactgta teetttaget toototagat<br>cagcgacccc tegtcacaat aaagataggg ggccagctga aggaggccct tetagaccag 34  | 80          |
| cagcgacccc tegteacaat aaagataggg ggetagetga aggregate tgeccetece 34 geeetggate cagatetget gtgeetteta gttgecagee atetgttgtt tgeccetece 34  | 340         |
| gccctggatc cagatctgct gtgccttcta gttgccagct dd-ggcgtaa taaaatgagg 35 ccgtgccttc cttgaccctg gaaggtgcca ctcccactgt cctttcctaa taaaatgagg 35   | 50 <b>0</b> |
| ccgtgccttc cttgaccctg gaaggtgcca ctcccactgc boodgagggggggggggggagg 36 aaattgcatc gcattgtctg agtaggtgtc attctattct   | 560         |
| aaattgcatc gcattgtctg agtaggtgtc attctattct   | 720         |
| acagcaaggg ggaggattgg gaagacaata gcaggcatge ggccagaaag aagcaggcac 3 tgggtaccca ggtgctgaag aattgacccg gttcctcctg ggccagaaag aagcaggcac 3   | 780         |
| tgggtaccca ggtgctgaag aattgacccg gttctttttg ggctaaa gccccactca 3' atccccttet etgtgacaca ccctgtccac gcccctggtt cttagttcca gccccactca 3'  | 840         |
| atccccttct ctgtgacaca ccctgtccac gcttctggtc coursels gtacttggag 3 taggacactc atagctcagg agggctccgc cttcaatccc acccgctaaa gtacttggag 3   | 900         |
| taggacacto atagotoagg agggotocgo ottoaacee abelya gtgggaagaa 3 cggtototoc otocotoato agocoaceaa accaaaceta gcotocaaga gtgggaagaa 3 cggtotocoa catgtgagga 3  | 960         |
| eggtetetee eteceteate ageceaceaa accadaceed solutions solutions ageteated ageceaceaa atgeetecaa catgtgagga 3 attaaageaa gataggetat taagtgeaga gggagagaaa atgeetecaa catgtgagga 3 attaaageaa gataggetat taagtgeaga ggataateg   | 020         |
| attaaagcaa gataggctat taagtgcaga gggagagada and attaatcttc cgcttcctcg 4 agtaatgaga gaaatcatag aattttaagg ccatcatggc cttaatcttc cgcttcctcg 4   | 080         |
| agtaatgaga gaaatcatag aattttaagg ctattatgge oobaa ctactcagc tcactcaaag 4 ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag 4  | 140         |
| ctcactgact cgctgcgctc ggtcgttcgg tcgcggcgag 535 5 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6   | 200         |
| gcggtaatac ggttatccac agaatcaggg gataatgcag gamas<br>ggccagcaaa aggccaggaa ccgtaaaaaag gccgcgttgc tggcgttttt ccataggctc 4   |             |
| - ·   |             |

```
egececetg acgageatea caaaaatega egeteaagte agaggtggeg aaaccegaca 4260
ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccq 4320
accetgeege ttaceggata cetgteegee ttteteeett egggaagegt ggegetteet 4380
catageteae getgtaggta teteagtteg gtgtaggteg ttegeteeaa getgggetgt 4440
gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag 4500
tccaaccegg taagacacga ettategeca etggcageag ccaetggtaa caggattage 4560
agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac 4620
actagaagaa cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga 4680
gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc 4740
aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg 4800
gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca 4860
aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt 4920
atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca 4980
gegatetgte tatttegtte atecatagtt geetgaeteg ggggggggg gegetgaggt 5040
ctgcctcgtg aagaaggtgt tgctgactca taccaggcct gaatcgcccc atcatccagc 5100
cagaaagtga gggagccacg gttgatgaga gctttgttgt aggtggacca gttggtgatt 5160
ttgaactttt gctttgccac ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc 5220
ttcaactcag caaaagttcg atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa 5280
tgctctgcca gtgttacaac caattaacca attctgatta gaaaaactca tcgagcatca 5340
aatgaaactg caatttattc atatcaggat tatcaatacc atatttttga aaaagccgtt 5400
tetgtaatga aggagaaaac teacegagge agtteeatag gatggcaaga teetggtate 5460
ggtctgcgat tccgactcgt ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa 5520
taaggttatc aagtgagaaa tcaccatgag tgacgactga atccggtgag aatggcaaaa 5580
gethatgeat thethteeag aethgiteaa caggecagee attacgeteg teatcaaaat 5640
cactegeate aaccaaaceg ttatteatte gtgattgege etgagegaga egaaataege 5700
gatcgctgtt aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc aggaacactg 5760
ccagcgcatc aacaatattt tcacctgaat caggatattc ttctaatacc tggaatgctg 5820
ttttcccggg gatcgcagtg gtgagtaacc atgcatcatc aggagtacgg ataaaatgct 5880
tgatggtcgg aagaggcata aattccgtca gccagtttag tctgaccatc tcatctgtaa 5940
catcattggc aacgctacct ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc 6000
catacaatcg atagattgtc gcacctgatt gcccgacatt atcgcgagcc catttatacc 6060
catataaatc agcatccatg ttggaattta atcgcggcct cgagcaagac gtttcccgtt 6120
gaatatggct cataacaccc cttgtattac tgtttatgta agcagacagt tttattgttc 6180
atgatgatat attittatct tgtgcaatgt aacatcagag attittgagac acaacgtggc 6240
tttccccccc cccccattat tgaagcattt atcagggtta ttgtctcatg agcggataca 6300
tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 6360
tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 6420
tcacgaggcc ctttcgtc
                                                                  6438
<210> 158
```

```
<211> 6486
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing SIV genes

<400> 158
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggggg tcagcggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
```

| tccaacatta | ccgccatgtt  | gacattgatt                              | attgactagt | tattaatagt | aatcaattac | 360 |
|------------|-------------|---|------------|------------|------------|-----|
|            | gttcatagcc. |   |            |            |            |     |
|            | tgaccgccca  |   |            |            |            |     |
|            | ccaataggga  |   |            |            |            |     |
|            | gcagtacatc  |   |            |            |            |     |
|            | tggcccgcct  |   |            |            |            |     |
|            | atctacgtat  |   |            |            |            |     |
|            | cgtggatagc  |   |            |            |            |     |
|            | agtttgtttt  |   |            |            |            |     |
|            | ttgacgcaaa  |   |            |            |            |     |
|            | gtgaaccgtc  |   |            |            |            |     |
|            | cgggaccgat  |   |            |            |            |     |
|            | aagagtgacg  |   |            |            |            |     |
|            | gctatactgt  |   |            |            |            |     |
|            | tatagettag  |   |            |            |            |     |
|            | gatactttcc  |   |            |            |            |     |
|            | atgccaatac  |   |            |            |            |     |
|            | ccatttatta  |   |            |            |            |     |
|            | attaaacata  |   |            |            |            |     |
|            | tctccggtag  |   |            |            |            |     |
|            | ggtcgctcgg  |   |            |            |            |     |
|            | ccaccaccac  |   |            |            |            |     |
|            | gtggagattg  |   |            |            |            |     |
|            | atgcaggcag  |   |            |            |            |     |
|            | tgttaacggt  |   |            |            |            |     |
|            | gacataatag  |   |            |            |            |     |
|            | gtcgtcgacg  |   |            |            |            |     |
|            | gagctggaga  |   |            |            |            |     |
|            | gtgtgggccg  |   |            |            |            |     |
|            | gagggctgcc  |   |            |            |            |     |
|            | ctgaagagcc  |   |            |            |            |     |
|            | aagcacaccg  |   |            |            |            |     |
|            | accgagacca  |   |            |            |            |     |
|            | tacceggtge  |   |            |            |            |     |
|            | aacgcetggg  |   |            |            |            |     |
|            | caggccctga  |   |            |            |            |     |
|            | gaccaccagg  |   |            |            |            |     |
|            | gacctgcagc  |   |            |            |            |     |
|            | gacategeeg  |   |            |            |            |     |
|            | aacccgatcc  |   |            |            |            |     |
|            | gtgcgcatgt  |   |            |            |            |     |
|            | cagagetacg  |   |            |            |            |     |
|            | aagaactgga  |   |            |            |            |     |
|            | ctgaagggcc  |   |            |            |            |     |
|            | ggcggcccgg  |   |            |            |            |     |
| _          | gtgcccatcc  |   | -          |            |            |     |
|            | aactgcggca  |   |            |            |            |     |
|            | tggaagtgcg  |   |            |            |            |     |
|            | ttaggccttg  |   |            |            |            |     |
|            | gggctgatgc  |   |            |            |            |     |
|            | cagttgggca  |   |            |            |            |     |
|            | gtgacagagg  |   |            |            |            |     |
|            | cgatcgaggg  |   |            |            |            |     |
|            | gccttctagt  |   |            |            |            |     |
|            | aggtgccact  |   |            |            |            |     |
|            | ,           | , |            | J. J       | 5          |     |

| attotctoao  | taggtgtgat   | tctattctgg | agaatagaat   | aaaaceaae  | 2002200000 | 2660 |
|-------------|--------------|------------|--------------|------------|------------|------|
| aggattggga  | agacaatagc   | aggcatgctg | aggatgcagt   | ggggcaggac | agcaaggggg | 3720 |
| tgctgaagaa  | ttgacccggt   | tcctcctggg | CCadaaadaa   | gggccccacg | ccccttctct | 3720 |
| gtgacacacc  | ctatccacac   | ccctggttct | tagttccacc   | ccactcata  | ggagagtgat | 3040 |
| ageteaggag  | gactecaect   | tcaatcccac | ccactaaaat   | acttogagg  | ggacactcat | 3040 |
| ccctcatcag  | ccaccaaac    | caaacctagc | ctccaagagt   | acceggageg | bassassas  | 3900 |
| taggetatta  | antacadada   | gagagaaaat | acctccaagage | tatasaasa  | taaagcaaya | 3900 |
| aatcatagaa  | ttttaacacc   | atcatggcct | taatettee    | chtactacat | caatgagaga | 4020 |
| ctacactagaa | teatteaact   | accacggccc | atatasaata   | ccccccgcc  | cactgacteg | 4080 |
| thatccacac  | aatcagggg    | gcggcgagcg | graceagere   | acccaaagge | ggtaatacgg | 4140 |
| accacacag   | attenangggga | taacgcagga | aagaacacgc   | gagcaaaagg | ccagcaaaag | 4200 |
| gecaygaace  | graaaaaggc   | cgcgttgctg | gegeeeeee    | ataggeteeg | ccccctgac  | 4260 |
| taggaccaca  | ttaaaategaeg | ctcaagtcag | aggeggegaa   | accegacagg | actataaaga | 4320 |
| caccaggege  | Liceceetgg   | aagctccctc | gracactere   | ctgttccgac | cctgccgctt | 4380 |
| accggatacc  | tgteegeett   | tctcccttcg | ggaagcgtgg   | cgctttctca | tageteaege | 4440 |
| tgtaggtatc  | reagrreggr   | gtaggtcgtt | cgctccaagc   | tgggctgtgt | gcacgaaccc | 4500 |
| cccgttcage  | ccgaccgctg   | cgccttatcc | ggtaactatc   | gtcttgagtc | caacccggta | 4560 |
| agacacgact  | tategeeact   | ggcagcagcc | actggtaaca   | ggattagcag | agcgaggtat | 4620 |
| graggcggtg  | ctacagagtt   | cttgaagtgg | tggcctaact   | acggctacac | tagaagaaca | 4680 |
| gtatttggta  | tctgcgctct   | gctgaagcca | gttaccttcg   | gaaaaagagt | tggtagctct | 4740 |
| tgatccggca  | aacaaaccac   | cgctggtagc | ggtggttttt   | ttgtttgcaa | gcagcagatt | 4800 |
| acgcgcagaa  | aaaaaggatc   | tcaagaagat | cctttgatct   | tttctacggg | gtctgacgct | 4860 |
| cagtggaacg  | aaaactcacg   | ttaagggatt | ttggtcatga   | gattatcaaa | aaggatcttc | 4920 |
| acctagatcc  | ttttaaatta   | aaaatgaagt | tttaaatcaa   | tctaaagtat | atatgagtaa | 4980 |
| acttggtctg  | acagttacca   | atgcttaatc | agtgaggcac   | ctatctcagc | gatctgtcta | 5040 |
| tttcgttcat  | ccatagttgc   | ctgactcggg | 9999999ggc   | gctgaggtct | gcctcgtgaa | 5100 |
| gaaggtgttg  | ctgactcata   | ccaggcctga | atcgccccat   | catccagcca | gaaagtgagg | 5160 |
| gagccacggt  | tgatgagagc   | tttgttgtag | gtggaccagt   | tggtgatttt | gaacttttgc | 5220 |
| tttgccacgg  | aacggtctgc   | gttgtcggga | agatgcgtga   | tctgatcctt | caactcagca | 5280 |
| aaagttcgat  | ttattcaaca   | aagccgccgt | cccgtcaagt   | cagcgtaatg | ctctgccagt | 5340 |
| gttacaacca  | attaaccaat   | tctgattaga | aaaactcatc   | gagcatcaaa | tgaaactgca | 5400 |
| atttattcat  | atcaggatta   | tcaataccat | atttttgaaa   | aagccgtttc | tgtaatgaag | 5460 |
| gagaaaactc  | accgaggcag   | ttccatagga | tggcaagatc   | ctggtatcgg | tetgegatte | 5520 |
| cgactcgtcc  | aacatcaata   | caacctatta | atttcccctc   | gtcaaaaata | aggttatcaa | 5580 |
| gtgagaaatc  | accatgagtg   | acgactgaat | ccggtgagaa   | tggcaaaagc | ttatgcattt | 5640 |
| ctttccagac  | ttgttcaaca   | ggccagccat | tacgctcgtc   | atcaaaatca | ctcgcatcaa | 5700 |
| ccaaaccgtt  | attcattcgt   | gattgcgcct | gagcgagacg   | aaatacgcga | togotottaa | 5760 |
| aaggacaatt  | acaaacagga   | atcgaatgca | accggcgcag   | gaacactgcc | agcgcatcaa | 5820 |
| caatattttc  | acctgaatca   | ggatattctt | ctaatacctg   | gaatgctgtt | ttcccgggga | 5880 |
| tcgcagtggt  | gagtaaccat   | gcatcatcag | gagtacqqat   | aaaatgcttg | atggtcggaa | 5940 |
| gaggcataaa  | ttccgtcagc   | cagtttagtc | tgaccatctc   | atctgtaaca | tcattggcaa | 6000 |
| cgctaccttt  | gccatgtttc   | agaaacaact | ctggcgcatc   | gggcttccca | tacaatcgat | 6060 |
| agattgtcgc  | acctgattgc   | ccgacattat | cqcqaqccca   | tttataccca | tataaatcag | 6120 |
| catccatgtt  | ggaatttaat   | cgcggcctcg | agcaagacgt   | ttcccqttqa | atatquetea | 6180 |
| taacacccct  | tgtattactq   | tttatgtaag | cagacagttt   | tattettcat | gatgatatat | 6240 |
| ttttatcttq  | tgcaatgtaa   | catcagagat | tttgagacac   | aacqtqqctt | tececece   | 6300 |
| cccattattq  | aagcatttat   | cagggttatt | qtctcatqaq   | cogatacata | tttgaatota | 6360 |
| tttagaaaaa  | taaacaaata   | ggggttccgc | gcacatttcc   | ccgaaaagto | ccacchaaca | 6420 |
| tctaaqaaac  | cattattatc   | atgacattaa | cctataaaaa   | taggcgtate | acqaqqqqq  | 6480 |
| ttcgtc      |              |            |              | -22202000  |            | 6486 |
| <b>J</b>    |              |            |              |            |            | 0400 |

<210> 159

<211> 9189

<212> DNA

<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 159 tegegegtht eggtgatgac ggtgaaaacc tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tqacqqtaaa tqqcccqcct qqcattatqc ccaqtacatg accttatqqq actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaaccgte agategeetg gagaegeeat ceaegetgtt ttgaccteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eeeetttgge 1080 tettatgeat getataetgt tittggetig gggeetatae acceegeti eettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct teteoggtag cggcggaget tecacatecg agecetggte ccatgcetec 1500 ageggeteat ggtegetegg cageteettg etectaaeag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgaqc gtgqagattq ggctcqcacq gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegecacea gacataatag etgacagaet aacagaetgt teettteeat gggtetttte 1860 tgcagtcacc gtcgtcgaca tgggcgcccg cgccagcgtg ctgagcggcg gcgagctgga 1920 ccgctgggag aagatccgcc tgcgccccgg cggcaagaag aagtacaagc tgaagcacat 1980 cgtgtgggcc agccgcgagc tggagcgctt cgccgtgaac cccggcctgc tggagaccag 2040 cgagggctgc cgccagatcc tgggccagct gcagcccagc ctgcagaccg gcagcgagga 2100 gctgcgcagc ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgagat 2160 caaggacacc aaggaggccc tggacaagat cgaggaggag cagaacaaga gcaagaagaa 2220 ggcccagcag gccgccgccg acaccggcca cagcaaccag gtgagccaga actacccat 2280 egtgeagaac atecagggee agatggtgea ceaggeeate ageceeegea ceetgaacge 2340 ctgggtgaag gtggtggagg agaaggcett cageeecgag gtgateeeca tgtteagege 2400 cctgagcgag ggcgccaccc cccaggacct gaacaccatg ctgaacaccg tgggcggcca 2460 ccaggccgcc atgcagatgc tgaaggagac catcaacgag gaggccgccg agtgggaccg 2520 cgtgcacccc gtgcacgccg gccccatcgc ccccggccag atgcgcgagc cccgcggcag 2580 cgacatcgcc ggcaccacca gcaccctgca ggagcagatc ggctggatga ccaacaaccc 2640 ccccatcccc gtgggcgaga tctacaagcg ctggatcatc ctgggcctga acaagatcgt 2700 gegeatgtac agececacca geatectgga cateegeeag ggeeecaagg agecetteeg 2760 cgactacgtg gaccgcttct acaagaccet gegegeegag caggecagec aggaggtgaa 2820 quactggatg accquagacco tgctggtgca gaacgccaac cccgactgca agaccatcct 2880 gaaggeeetg ggeeegeeg ceaccetgga ggagatgatg accgeetgee agggegtggg 2940

|              | 1                        |            |             |             |            |      |
|--------------|--------------------------|------------|-------------|-------------|------------|------|
| cggccccggc   | cacaaggccc               | gcgtgctggc | cgaggccatg  | agccaggtga  | ccaacagcgc | 3000 |
| caccatcatg   | atgcagcgcg               | gcaacttccg | caaccagcgc  | aagatcgtga  | agtgcttcaa | 3060 |
| ctgcggcaag   | gagggccaca               | ccgcccgcaa | ctgccgcgcc  | ccccgcaaga  | agggctgctg | 3120 |
| gaagtgcggc   | aaggagggcc               | accagatgaa | ggactgcacc  | gagcgacagg  | ctaataggga | 3180 |
| agatctggcc   | ttcccacaag               | ggaaggccag | ggaattttct  | tcagagcaga  | ccagagccaa | 3240 |
| cagccccacc   | agaagagagc               | ttcaggtttg | gggaagagac  | aacaactccc  | tctcagaagc | 3300 |
| aggagccgat   | agacaaggaa               | ctgtatcctt | tagcttccct  | cagatcactc  | tttggcagcg | 3360 |
| acccctcgtc   | acaataaaga               | tagggggcca | gctgaaggag  | gcccttctag  | acaccggcgc | 3420 |
| cgacgacacc   | gtgctggagg               | agatgaacct | gcccggccgc  | tggaagccca  | agatgatcgg | 3480 |
| cggcatcggc   | ggcttcatca               | aggtgggcca | gtacgaccag  | atcctgatcg  | agatctgcgg | 3540 |
| ccacaaggcc   | atcggcaccg               | tgctggtggg | ccccaccccc  | gtgaacatca  | tcggccgcaa | 3600 |
| cctgctgacc   | cagatcggct               | gcaccctgaa | cttccccatc  | agccccatcg  | agaccgtgcc | 3660 |
| cgtgaagctg   | aagcccggca               | tggacggccc | caaggtgaag  | cagtggcccc  | tgaccgagga | 3720 |
| gaagatcaag   | gccctggtgg               | agatctgcac | cgagatggag  | aaggagggca  | agatcagcaa | 3780 |
| gatcggcccc   | gagaacccct               | acaacacccc | cgtgttcgcc  | atcaagaaga  | aggacagcac | 3840 |
| caagtggcgc   | aagctggtgg               | acttccgcga | gctgaacaag  | cgcacccagg  | acttctggga | 3900 |
| ggtgcagctg   | ggcatccccc               | accccgccgg | cctgaagcag  | aagaagagcg  | tgaccgtgct | 3960 |
| ggacgtgggc   | gacgcctact               | tcagcgtgcc | cctggacaag  | gacttccgca  | agtacaccgc | 4020 |
| cttcaccatc   | cccagcatca               | acaacgagac | ccccggcatc  | cgctaccagt  | acaacgtgct | 4080 |
| gccccagggc   | tggaagggca               | gccccgccat | cttccagtgc  | agcatgacca  | agatcctgga | 4140 |
| gcccttccgc   | aagcagaacc               | ccgacatcgt | gatctaccag  | tacatggacg  | acctgtacgt | 4200 |
| gggcagcgac   | ctggagatcg               | gccagcaccg | caccaagatc  | gaggagctgc  | gccagcacct | 4260 |
| gctgcgctgg   | ggcttcacca               | ccccgacaa  | gaagcaccag  | aaggagcccc  | ccttcctgtg | 4320 |
| gatgggctac   | gagctgcacc               | ccgacaagtg | gaccgtgcag  | cccatcgtgc  | tgcccgagaa | 4380 |
| ggacagctgg   | accgtgaacg               | acatccagaa | gctggtgggc  | aagctgaact  | gggccagcca | 4440 |
| gatctacgcc   | ggcatcaagg               | tgcgccagct | gtgcaagctg  | ctgcgcggca  | ccaaggccct | 4500 |
| gaccgaggtg   | gtgcccctga               | ccgaggaggc | cgagctggag  | ctggccgaga  | accgcgagat | 4560 |
| cctgaaggag   | cccgtgcacg               | gcgtgtacta | cgaccccagc  | aaggacctga  | tegeegagat | 4620 |
| ccagaagcag   | ggccagggcc               | agtggaccta | ccagatctac  | caggagccct  | tcaagaacct | 4680 |
| gaagaccggc   | aagtacgccc               | gcatgaaggg | cgcccacacc  | aacgacgtga  | agcagctgac | 4740 |
| cgaggccgtg   | cagaagatcg               | ccaccgagag | catcgtgatc  | tggggcaaga  | ccccaagtt  | 4800 |
| caagetgeee   | atccagaagg               | agacctggga | ggcctggtgg  | accgagtact  | ggcaggccac | 4860 |
| ctggatecee   | gagtgggagt               | tcgtgaacac | ccccccctg   | gtgaagctgt  | ggtaccagct | 4920 |
| ggagaaggag   | cccatcatcg               | gcgccgagac | cttctacgtg  | gacggcgccg  | ccaaccgcga | 4980 |
| gaccaagctg   | ggcaaggccg               | gctacgtgac | cgaccgcggc  | cgccagaagg  | tggtgcccct | 5040 |
| gaccgacacc   | accaaccaga               | agaccgagct | gcaggccatc  | cacctggccc  | tgcaggacag | 5100 |
| cggcctggag   | gtgaacatcg               | tgaccgacag | ccagtacgcc  | ctgggcatca  | tccaggccca | 5160 |
| gcccgacaag   | agcgagagcg               | agctggtgag | ccagatcatc  | gagcagctga  | tcaagaagga | 5220 |
| gaaggtgtac   | ctggcctggg               | tgcccgccca | caagggcatc  | ggcggcaacg  | agcaggtgga | 5280 |
| eggeetggtg   | agegeeggea               | teegeaaggt | getgtteetg  | gacggcatcg  | acaaggccca | 5340 |
| ggaggagcac   | gagaagtacc               | acagcaactg | gegegeeatg  | gccagcgact  | tcaacctgcc | 5400 |
| cetegeggeg   | gccaaggaga               | cegtggeeag | cigogacaag  | tgccagctga  | agggcgaggc | 5460 |
| gaggaagge    | caggtggact               | geageeeegg | catctggcag  | ctggactgca  | cccacctgga | 5520 |
| gggcaaggcg   | atcctggtgg               | ccgcgcacgc | ggccagcgge  | cacaccgagg  | ccgaggtgat | 5580 |
| catasaasaa   | accggccagg               | agacegeeta | assabbases  | aagetggeeg  | gccgctggcc | 5640 |
| ctactaataa   | gtgcacaccg               | acaacggcag | caacticace  | agcaccaccg  | tgaaggccgc | 5700 |
| catastasa    | gccggcatca               | agcaggagcc | cygcatecee  | cacaaccccc  | agagccaggg | 5760 |
| caaccacctc   | agcatgaaca               | taceastace | gaayattatt  | ggccaggtgc  | gegaceagge | 5820 |
| cadcatcacc   | aagaccgccg               | ccaccasaca | categicate  | cacaacccca  | agegeaaggg | 2880 |
| ascessage    | ggctacagcg               | acatoscos  | catcytygac  | accaucgeca  | cegacateca | 5940 |
| caccoccase   | ctgcagaagc               | agaccaccaa | yacccayaac  | bassasses   | actaccgcga | 6000 |
| agtastorsa   | cccgtgtgga<br>gacaacagcg | agggccccgc | agtaccase=  | Lygaagggcg  | agggegeegt | 6060 |
| cgactaccag   | aagcagatgg               | ccaccaagge | ctacatacac  | agecagaaa   | agaccatccg | 0120 |
| . ggaattetge | tgtgccttct               | aattaccaac | catctottot  | ttaccactage | cccatacatt | 0700 |
| JJ=======    | - 5 - 5 - 5 - 5 - 5 - 5  |            | -accondecat | cogocococo  | caagagaaat | 0240 |

| ccttgaccct g | gaaggtgcc  | actcccactq  | tectttecta | ataaaatgag | gaaattgcat | 6300 |
|--------------|------------|-------------|------------|------------|------------|------|
| cgcattgtct g | agtaggtgt  | cattctattc  | taaagagtaa | ggtgggcag  | gacagcaagg | 6360 |
| gggaggattg g | gaagacaat  | agcaggcatg  | ctggggatgc | ggtgggctct | atgggtaccc | 6420 |
| aggtgctgaa g | aattgaccc  | ggttcctcct  | qqqccagaaa | gaagcaggca | catccccttc | 6480 |
| tctgtgacac a | ccctgtcca  | cacccctaat  | tcttagttcc | agececaete | ataggacact | 6540 |
| catageteag g | aagactcca  | ccttcaatcc  | cacccactaa | agtacttgga | gcggtctctc | 6600 |
| cctccctcat c | rageceacea | aaccaaacct  | agectecaag | agtaggaaga | aattaaagca | 6660 |
| agataggcta t | taantocan  | acccacacaa  | aatocctcca | acatotoago | aagtaatgag | 6720 |
| agaaatcata g | raattttaag | accatcataa  | ccttaatctt | ccacttcctc | gctcactgac | 6780 |
| tegetgeget e | rateattea  | actacaacaa  | acaatatcaa | ctcactcaaa | ggcggtaata | 6840 |
| cggttatcca c | ragaatcaca | geegeggege  | gaggagaga  | tataaacaaa | aggccagcaa | 6900 |
| aaggccagga a | agaaccagg  | ggacacogta  | ctaacatttt | tocatagot  | ccaccccct  | 6960 |
| gacgagcatc a | ccycaaaaa  | aggegegege  | cagacataca | daaacccdac | aggactataa | 7020 |
| agataccagg o | caaaaaccg  | tessasstes  | ctcatacact | ctcctattcc | aggactacca | 7080 |
| cttaccggat a | geeeeee    | attatata    | tragageger | tagcactttc | tratagetra | 7140 |
| cttaccggat a | accegecege | ectitettett | attactes   | agetagaeta | tatacacaaa | 7200 |
| cgctgtaggt a | acccagccc  | ggtgtaggtt  | teeestaast | ageceggeeg | atcasacca  | 7260 |
| cccccgttc a  | agecegaceg | etgegeetta  | accaptant  | accyccicya | gaccaacccg | 7200 |
| gtaagacacg a | acttategee | actggcagca  | gecaetggta | acaggactag | cagagegagg | 7320 |
| tatgtaggcg g | gtgctacaga | gttettgaag  | tggtggteta | actacggeta | cactagaaga | 7360 |
| acagtatttg   | gtatetgege | tctgctgaag  | ccagttacct | teggaaaaag | ageeggeage | 7440 |
| tcttgatccg   | gcaaacaaac | caccgctggt  | agcggtggtt | tttttgtttg | caagcagcag | 7500 |
| attacgcgca g | gaaaaaaagg | atctcaagaa  | gateetttga | tetttetae  | ggggtctgac | 7560 |
| gctcagtgga a | acgaaaactc | acgttaaggg  | attttggtca | tgagattatc | aaaaaggatc | 7620 |
| ttcacctaga t | tccttttaaa | ttaaaaatga  | agttttaaat | caatctaaag | tatatatgag | 7680 |
| taaacttggt ( | ctgacagtta | ccaatgctta  | atcagtgagg | cacctatctc | agcgatctgt | 7740 |
| ctatttcgtt ( | catccatagt | tgcctgactc  | aaaaaaaaaa | ggcgctgagg | tetgeetegt | 7800 |
| gaagaaggtg ( | ttgctgactc | ataccaggcc  | tgaatcgccc | catcatccag | ccagaaagtg | 7860 |
| agggagccac   | ggttgatgag | agctttgttg  | taggtggacc | agttggtgat | tttgaacttt | 7920 |
| tgctttgcca   | cggaacggtc | tgcgttgtcg  | ggaagatgcg | tgatctgatc | cttcaactca | 7980 |
| gcaaaagttc g | gatttattca | acaaagccgc  | cgtcccgtca | agtcagcgta | atgctctgcc | 8040 |
| agtgttacaa   | ccaattaacc | aattctgatt  | agaaaaactc | atcgagcatc | aaatgaaact | 8100 |
| gcaatttatt   | catatcagga | ttatcaatac  | catatttttg | aaaaagccgt | ttctgtaatg | 8160 |
| aaggagaaaa   | ctcaccgagg | cagttccata  | ggatggcaag | atcctggtat | cggtctgcga | 8220 |
| ttccgactcg   | tccaacatca | atacaaccta  | ttaatttccc | ctcgtcaaaa | ataaggttat | 8280 |
| caagtgagaa   | atcaccatga | gtgacgactg  | aatccggtga | gaatggcaaa | agcttatgca | 8340 |
| tttctttcca   | gacttgttca | acaggccagc  | cattacgctc | gtcatcaaaa | tcactcgcat | 8400 |
| caaccaaacc   | gttattcatt | cgtgattgcg  | cctgagcgag | acgaaatacg | cgatcgctgt | 8460 |
| taaaaggaca   | attacaaaca | ggaatcgaat  | gcaaccggcg | caggaacact | gccagcgcat | 8520 |
| caacaatatt   | ttcacctgaa | tcaggatatt  | cttctaatac | ctggaatgct | gttttcccgg | 8580 |
| ggatcgcagt   | ggtgagtaac | catgcatcat  | caggagtacg | gataaaatgc | ttgatggtcg | 8640 |
| gaagaggcat   | aaattccgtc | agccagttta  | gtctgaccat | ctcatctgta | acatcattgg | 8700 |
| caacgctacc   | tttgccatgt | ttcagaaaca  | actctggcgc | atcgggcttc | ccatacaatc | 8760 |
| gatagattgt   | cgcacctgat | tgcccgacat  | tatcgcgagc | ccatttatac | ccatataaat | 8820 |
| cagcatccat   | gttggaattt | aatcgcggcc  | tcgagcaaga | cgtttcccgt | tgaatatggc | 8880 |
| tcataacacc   | ccttgtatta | ctgtttatgt  | aagcagacag | ttttattgtt | catgatgata | 8940 |
| tatttttatc   | ttgtgcaatg | taacatcaga  | gattttgaga | cacaacgtgg | ctttccccc  | 9000 |
| cccccatta    | ttgaagcatt | tatcagggtt  | attgtctcat | gagcggatac | atatttgaat | 9060 |
| gtatttagaa   | aaataaacaa | ataggggttc  | cgcgcacatt | tccccgaaaa | gtgccacctg | 9120 |
| acgtctaaga   | aaccattatt | atcatgacat  | taacctataa | aaataggcgt | atcacgaggc |      |
| cctttcgtc    |            |             |            |            |            | 9189 |
|              |            |             |            |            |            |      |

<210> 160 <211> 9285

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing SIV genes

<400> 160

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tocaacatta cogocatgit gacattgatt altgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaqgtct atataaqcag 900 agctegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac egggacegat ceagecteeg eggeegggaa eggtgeattg gaacgeggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getataetgt ttttggettg gggeetatae accecegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggactetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tqqaqqecaq aettaqqeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegecacca gacataatag etgacagact aacagactgt teethtecat gggteththe 1860 tgcagtcacc gtcgtcgacg ccgccaccat gggcgtgcgc aacagcgtgc tgagcggcaa 1920 gaaggccgac gagctggaga agatccgcct gcgcccgaac ggcaagaaga agtacatgct 1980 gaagcacgtg gtgtgggccg ccaacgagct ggaccgcttc ggcctggccg agagcctgct 2040 ggagaacaag gagggctgcc agaagatcct gagcgtgctg gcgccgctgg tgcccaccgg 2100 cagegagaac etgaagagee tgtacaacac egtgtgegtg atetggtgea tecaegeega 2160 ggagaaggtg aagcacaccg aggaggccaa gcagatcgtg cagcgccacc tggtggtgga 2220 gaceggeace accgagacea tgeccaagac cageegeeeg accgegeeca geageggaeg 2280 cggcggcaac tacccggtgc agcagatcgg cggcaactac gtgcacctgc cgctgagccc 2340 gcgcaccctg aacgcctggg tgaagctgat cgaggagaag aagttcggcg ccgaggtggt 2400 geceggette caggeectga gegagggetg caegeectae gacateaace agatgetgaa 2460 ctgcgtgggc gaccaccagg ccgccatgca gatcatccgc gacatcatca acgaggaggc 2520 egeogactgg gacetgeage accegeagee egegeogeag cagggecage tgegegagee 2580 cageggcage gacategeeg geaceaceag cagegtggae gageagatee agtggatgta 2640 ccgccagcag aacccgatcc cggtgggcaa catctaccgc cgctggatac agctgggcct 2700 qcagaagtgc gtgcgcatgt acaacccgac caacatcctg gacgtgaagc agggcccgaa 2760 ggagccettc cagagctacg tggaccgctt ctacaagagc ctgcgcgccg agcagaccga 2820 cgccgccgtg aagaactgga tgacccagac cctgctgatc cagaacgcca acccggactg 2880

| caagetagta  | ctassagacc   | tagacataaa   | cccgaccctg   | gaggagatgc              | tgaccgcctg | 2940 |
|-------------|--------------|--------------|--------------|-------------------------|------------|------|
| ccagggggg   | aacaacccaa   | accagaagge   | ccacctaata   | gccgaggccc              | tgaaggaggc | 3000 |
| cctagggegeg | ggoggoobgg   | cattcacaac   | cacccaacaa   | cgcggcccgc              | gcaagcccat | 3060 |
| caachachaa  | aactocooca   | aggagggca    | cadcdcccdc   | cagtgccgcg              | egeegegeeg | 3120 |
| ccagagatas  | tagaagtaca   | gcaagatgga   | ccacqtqatq   | gccaagtgcc              | cggaccgcca | 3180 |
| ccagggctgc  | ttaggaugugug | atccatagaa   | aaadaadccd   | cgcaatttcc              | cgatggctca | 3240 |
| ggegggeeee  | aggetta      | caactoctcc   | accadaddac   | ccagctgtgg              | atctgctaaa | 3300 |
| agegeareag  | gggctgatge   | accaccacac   | adaaaadcad   | agagaaagca              | gagagaagcc | 3360 |
| gaactacatg  | cageegggea   | atttactaca   | cctcaattct   | ctctttggag              | gagaccagta | 3420 |
| ccacaaggag  | grgacagagg   | caraccata    | gaggtgcttc   | tagacaccgg              | caccaacaac | 3480 |
| gcactgcat   | cgatcgaggg   | actoracces   | cactacaccc   | cgaagatcgt              | agacaacatc | 3540 |
| agcatcgtga  | eeggeacega   | getgggeteg   | aacatagaga   | tegaggtget              | agacaaacac | 3600 |
| ggcggcttca  | teaacacegg   | agagtacaag   | cccatcaaca   | tettegacea              | caacctgctg | 3660 |
| atcaagggca  | ccatcatgac   | cggcgacacg   | atcaccaaca   | tetteggeeg              | gaagotggcc | 3720 |
| accgccctgg  | gcatgagcct   | gaacttcccc   | accyccaagg   | tggagcccgt              | gaagagatc  | 3780 |
| ctgaagcccg  | gcaaggacgg   | ceceaagetg   | aagcagragge  | ccctgagcaa              | adadaccca  | 3840 |
| gtggccctgc  | gcgagatctg   | cgagaagatg   | gagaaggacg   | gccagctgga              | gaaggeeeea | 3900 |
| cccaccaacc  | cctacaacac   | gcccaccttc   | gccatcaaga   | agaaggacaa              | gaacaagtgg | 3960 |
| cgcatgctga  | tcgacttccg   | cgagctgaac   | egegegaeee   | aggacttcac              | cgaggcgcag | 4020 |
| ctgggcatcc  | cgcaccccgc   | caaccraacc   | aagcgcaagc   | gcatcaccgt              | getggacate | 4020 |
| ggcgacgcct  | acttcagcat   | cccctggac    | gaggagttcc   | gccagtacac              | egeecceaec | 4140 |
| ctgcccagcg  | tgaacaacgc   | cgagcccggc   | aagcgctaca   | tctacaaggt              | gergeeedag | 4740 |
| ggctggaagg  | gcagccccgc   | catcttccag   | tacaccatgc   | gccacgtgct              | ggageeette | 4200 |
| cgcaaggcca  | accccgacgt   | gaccetggtg   | cagtacatgc   | atgacatcct              | gacegecage | 4200 |
| gaccgcaccg  | acctggagca   | cgaccgcgtg   | gtgctgcaga   | gcaaggagct              | geegaacage | 4320 |
| atcggcttca  | gcacgcccga   | ggagaagttc   | cagaaggacc   | caccetteca              | gtggatgggc | 4380 |
| tacgagctgt  | ggcccaccaa   | gtggaagctg   | cagaagatcg   | agctgcccca              | gcgcgagacc | 4440 |
| tggaccgtga  | acgacatcca   | gaagctggtg   | ggcgtgctga   | actgggccgc              | ccagatetac | 4500 |
| cccggcatca  | agaccaagca   | cctgtgccgc   | ctgatccgcg   | gcaagatgac              | cctgaccgag | 4560 |
| gaggtgcagt  | qqaccqaqat   | ggccgaggcc   | gagtacgagg   | agaacaagat              | catcctgagc | 4620 |
| caggagcagg  | agggctgcta   | ctaccaggag   | ggcaagcccc   | tggaggccac              | cgtgatcaag | 4680 |
| agccaggaca  | accagtggag   | ctacaagatc   | caccaggagg   | acaagatcct              | gaaggtgggc | 4/40 |
| aagttcgcca  | agatcaagaa   | cacccacacc   | aacggcgtgc   | gcctgctggc              | ccacgtgatc | 4800 |
| cagaagatcg  | qcaaqqaqqc   | catcgtgatc   | : tggggccagg | tgcccaagtt              | ccacctgccc | 4860 |
| atagagaagg  | acqtqtqqqa   | gcagtggtgg   | accgactact   | ggcaggtgac              | ctggataccc | 4920 |
| gagtoggact  | tcatcagcac   | ccctcccctg   | gtgcgcctgg   | tgttcaacct              | ggtgaaggac | 4980 |
| cccatcgagg  | qcqaqqaqac   | ctactacaco   | : gacggcagct | gcaacaagca              | gagcaaggag | 5040 |
| ggcaaggccg  | gctacatcac   | cgaccgcggc   | aaggacaagg   | tgaaggtgct              | ggagcagacc | 2100 |
| accaaccago  | aggccgagct   | cgaggccttc   | : ctgatggccc | : tgaccgacag            | cggccccaag | 2100 |
| accaacatca  | tcqtqqacag   | ccagtacgtg   | , atgggcatca | tcaccggctg              | ccccaccgag | 5220 |
| agcgagagcg  | gcctggtgaa   | ccagatcato   | : gaggagatga | , tcaagaagag            | cgagatctac | 5280 |
| ataacetaaa  | tacccaccca   | caagggcato   | : ggcggcaacc | : aggagatcga            | ccacctggtg | 5340 |
| agccagggca  | tccgccaggt   | getgtteetg   | gagaagatcg   | <sub>l</sub> agcccgccca | ggaggagcac | 5400 |
| gacaagtaco  | : acaqcaacgt | gaaggagctg   | , gtgttcaagt | : teggeetgee            | ccgcatcgtg | 5460 |
| acccaccaga  | tegtggacac   | ctgcgacaag   | , tgccaccaga | agggcgaggc              | catccacggc | 5520 |
| caggccaaca  | qcqacctggg   | cacctggcag   | g atggcctgca | a cccacctgga            | gggcaagatc | 5580 |
| atcatcatas  | ccatacacat   | ggccagcggg   | : ttcatcgagg | g ccgaggtgat            | cccgcaggag | 5640 |
| accooccoco  | : agaccgccct | gttactgata   | g aagctggccg | g geegetggee            | catcacccac | 5700 |
| ctocacacco  | acaacqqcqc   | caacttcgc    | agccaggagg   | g tgaagatggt            | ggcctggtgg | 5760 |
| accaacatco  | , agcacacctt | : cggcgtgccc | : tacaacccg  | agagccaggg              | cgtggtggag | 5820 |
| accatasacc  | accacctga    | qaaccaqat    | gaccgcatco   | gcgagcaggc              | caacagcgtg | 5880 |
| gagaccatco  | tactaataac   | cgtacactac   | atgaacttca   | agegeegegg              | cggcatcggc | 5940 |
| gagatoacog  | ccaccaaaca   | cctgatcaa    | c atgatcacca | ccgagcagga              | gatccagtto | 6000 |
| Ancarage    | agaacagcaa   | gttcaagaa    | c ttccacatal | t actaccgcga            | gggccgcgac | 6060 |
| cagcagage   | a agggccccac | cgagetget    | q tqqaaqqqc  | g agggcgccgt            | gatectgaag | 6120 |
| atacacacac  | r acatcaaoot | gataccca     | c cqcaaqqcc  | a agatcatcaa            | ggactacggo | 6180 |
| a caaacacci | 2 464664433  |              |              | -                       |            |      |

| ggcggcaagg | aggtggacag | cagcagccac | atggaggaca | ccggcgaggc | ccgcgaggtg | 6240 |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | ccttctagtt |      |
|            |            |            |            |            | ggtgccactc |      |
|            |            |            |            |            | aggtgtcatt |      |
|            |            |            |            |            | gacaatagca |      |
|            |            |            |            |            | tgacccggtt |      |
|            |            |            |            |            | tgtccacgcc |      |
|            |            |            |            |            | gctccgcctt |      |
|            |            |            |            |            | ccaccaaacc |      |
|            |            |            |            |            | gtgcagaggg |      |
|            |            |            |            |            | tttaaggcca |      |
|            |            |            |            |            | cgttcggctg |      |
|            |            |            |            |            | atcaggggat |      |
|            |            |            |            |            | taaaaaggcc |      |
|            |            |            |            |            | aaatcgacgc |      |
|            |            |            |            |            | tececetgga |      |
|            |            |            |            |            | gtccgccttt |      |
|            |            |            |            |            | cagttcggtg |      |
|            |            |            |            |            | cgaccgctgc |      |
|            |            |            |            |            | atcgccactg |      |
|            |            |            |            |            | tacagagttc |      |
|            |            |            |            |            | ctgcgctctg |      |
|            |            |            |            |            | acaaaccacc |      |
|            |            |            |            |            | aaaaggatct |      |
|            |            |            |            |            | aaactcacgt |      |
|            |            |            |            |            | tttaaattaa |      |
|            |            |            |            |            | cagttaccaa |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | catagttgcc |      |
|            |            |            |            |            | tgactcatac |      |
|            |            |            |            |            | gatgagaget |      |
|            |            |            |            |            | acggtctgcg |      |
|            |            |            |            |            | tattcaacaa |      |
|            |            |            |            |            | ttaaccaatt |      |
|            |            |            |            |            | tcaggattat |      |
|            |            |            |            |            | ccgaggcagt |      |
|            |            |            |            |            | acatcaatac |      |
|            |            |            |            |            | ccatgagtga |      |
|            |            |            |            |            | tgttcaacag |      |
|            |            |            |            |            | ttcattcgtg |      |
|            |            |            |            |            | caaacaggaa |      |
|            |            |            |            |            | cctgaatcag |      |
|            |            |            |            |            | agtaaccatg |      |
|            |            |            |            |            | tccgtcagcc |      |
|            |            |            |            |            | ccatgtttca |      |
|            |            |            |            |            | cctgattgcc |      |
|            |            |            |            |            | gaatttaatc |      |
|            |            |            |            |            | gtattactgt |      |
|            |            |            |            |            | gcaatgtaac |      |
|            |            |            |            |            | agcatttatc |      |
|            |            |            |            |            | aaacaaatag |      |
|            |            |            |            |            | attattatca |      |
| Lyacattaac | ccacaaaaat | aggegtatea | cgaggccctt | ccgcc      |            | 9285 |

<210> 161 <211> 7897

<212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cocqcetggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 etcegeecea ttgaegeaaa tgggeggtag gegtgtaegg tgggaggtet atataageag 900 agetegttta qtqaaccgtc agategeetg gagacgecat ccacgetqtt ttqaccteca 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae accceegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accaetecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ecatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca tgagggaaga tctggccttc ccacaaggga aggccaggga 1920 attitettea gageagacea gageeaacag eeccaceaga agagagette aggtitgggg 1980 aagagacaac aactccctct cagaagcagg agccgataga caaggaactg tatcctttag 2040 cttccctcag atcactcttt ggcagcgacc cctcgtcaca ataaagatag ggggccagct 2100 gaaggaggee ettetagaea eeggegeega egacacegtg etggaggaga tgaacetgee 2160 eggeegetgg aageecaaga tgateggegg categgegge tteateaagg tgggeeagta 2220 cgaccagatc ctgatcgaga tctgcggcca caaggccatc ggcaccgtgc tggtgggccc 2280 caccecegtg aacatcateg geogeaacet getgacecag ateggetgea eeetgaactt 2340 ccccatcagc cccatcgaga ccgtgcccgt gaagctgaag cccggcatgg acggccccaa 2400 ggtgaagcag tggcccctga ccgaggagaa gatcaaggcc ctggtggaga tctgcaccga 2460 gatggagaag gagggcaaga tcagcaagat cggccccgag aacccctaca acacccccgt 2520 gttcgccatc aagaagaagg acagcaccaa gtggcgcaag ctggtggact tccgcgagct 2580 gaacaagege acceaggact tetgggaggt geagetggge atceeceace eegeeggeet 2640 gaaqcagaag aaqaqcqtqa ccqtqctgga cgtgggcgac gcctacttca gcqtqcccct 2700 ggacaaggac ttccgcaagt acaccgcctt caccatcccc agcatcaaca acgagacccc 2760

| cggcatccgc | taccagtaca   | acgtgctgcc | ccagggctgg | aagggcagcc | ccgccatctt | 2820 |
|------------|--------------|------------|------------|------------|------------|------|
|            |              |            |            |            | acatcgtgat |      |
|            |              |            |            |            | agcaccgcac |      |
|            |              |            |            |            | ccgacaagaa |      |
|            |              |            |            |            | acaagtggac |      |
|            |              |            |            |            | tccagaagct |      |
|            |              |            |            |            | gccagctgtg |      |
|            |              |            |            |            | aggaggccga |      |
|            |              |            |            |            | tgtactacga |      |
|            |              |            |            |            | ggacctacca |      |
|            |              |            |            |            | tgaagggcgc |      |
|            |              |            |            |            | ccgagagcat |      |
|            |              |            |            |            | cctgggaggc |      |
|            |              |            |            |            | tgaacacccc |      |
|            |              |            |            |            | ccgagacctt |      |
|            |              |            |            |            | acgtgaccga |      |
|            |              |            |            |            | ccgagctgca |      |
|            |              |            |            |            | ccgacagcca |      |
|            |              |            |            |            | tggtgagcca |      |
|            |              |            |            |            | ccgcccacaa |      |
|            |              |            |            |            | gcaaggtgct |      |
|            |              |            |            |            | gcaactggcg |      |
|            |              |            |            |            | tggccagctg |      |
|            |              |            |            |            | gccccggcat |      |
|            |              |            |            |            |            |      |
|            |              |            |            |            | tgcacgtggc |      |
|            |              |            |            |            | ccgcctactt |      |
|            |              |            |            |            | acggcagcaa |      |
|            |              |            |            |            | aggagttcgg |      |
|            |              |            |            |            | agctgaagaa |      |
|            |              |            |            |            | agatggccgt |      |
|            |              |            |            |            | gcgagcgcat |      |
| cgrggacare | acceptedacce | acacccagac | caaggagetg | cagaagcaga | tcaccaagat | 4580 |
|            |              |            |            |            | gccccgccaa |      |
|            |              |            |            |            | tcaaggtggt |      |
|            |              |            |            |            | gcgacgactg |      |
|            |              |            |            |            | ttgccagcca |      |
|            |              |            |            |            | tcccactgtc |      |
|            |              |            |            |            | ttctattctg |      |
|            |              |            |            |            | caggcatgct |      |
|            |              |            |            |            | ttcctcctgg |      |
|            |              |            |            |            | ccctggttc  |      |
|            |              |            |            |            | ttcaatccca |      |
|            |              |            |            |            | ccaaacctag |      |
|            |              |            |            |            | ggagagaaaa |      |
|            |              |            |            |            | catcatggcc |      |
|            |              |            |            |            | tgcggcgagc |      |
|            |              |            |            |            | ataacgcagg |      |
|            |              |            |            |            | ccgcgttgct |      |
|            |              |            |            |            | gctcaagtca |      |
|            |              |            |            |            | gaagctccct |      |
|            |              |            |            |            | ttctcccttc |      |
|            |              |            |            |            | tgtaggtcgt |      |
| tcgctccaag | ctgggctgtg   | tgcacgaacc | ccccgttcag | cccgaccgct | gcgccttatc | 5940 |
|            |              |            |            |            | tggcagcagc |      |
| cactggtaac | aggattagca   | gagcgaggta | tgtaggcggt | gctacagagt | tcttgaagtg | 6060 |
|            |              |            |            |            |            |      |

```
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 6120
agttaccttc ggaaaaagag ttggtagctc ttgatccqqc aaacaaacca ccqctqqtaq 6180
eggtggtttt tttgtttgca agcagcagat tacgegcaga aaaaaaggat ctcaagaaga 6240
tectttgate ttttctaegg ggtetgaege teagtggaae gaaaacteae gttaagggat 6300
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 6360 -
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 6420
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttq cctqactcqq 6480
ggggggggg cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat accaggcctg 6540
aatogoocca toatocagoo agaaagtgag ggagocacgg ttgatgagag otttgttqta 6600
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcgqq 6660
aaqatqcqtg atctgatcct tcaactcaqc aaaaqttcqa tttattcaac aaaqccqccq 6720
tecegteaag teagegtaat getetgeeag tgttacaace aattaaceaa ttetgattag 6780
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 6840
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 6900
atggcaagat cetggtateg gtetgegatt cegactegte caacateaat acaacetatt 6960
aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 7020
teeggtgaga atggcaaaag ettatgcatt tettteeaga ettgtteaac aggceageca 7080
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 7140
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 7200
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 7260
tetaataect ggaatgetgt ttteeegggg ategeagtgg tgagtaacca tgeateatea 7320
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 7380
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 7440
tetggegeat egggetteec atacaatega tagattgteg cacetgattg eccgacatta 7500
tegegagece atttatacce atataaatea geatecatgt tggaatttaa tegeggeete 7560
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 7620
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 7680
ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta tcagggttat 7740
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 7800
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 7860
acctataaaa ataggcgtat cacgaggccc tttcgtc
                                                                  7897
<210> 162
<211> 7984
<212> DNA
<213> Artificial Sequence
<223> plasmid pVR1012x/s containing SIV genes
<400> 162
tegegegitt eggigatgae ggitgaaaacc tetgacacat geageteeeg gagaeggica 60 .
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tocaacatta cogocatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
```

catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720

| catcaatggg  | cgtggatagc | ggtttgactc | acggggattt | ccaaqtetee  | accccattos | 780  |
|-------------|------------|------------|------------|-------------|------------|------|
| cgtcaatggg  | agtttgttt  | qqcaccaaaa | tcaacgggac | tttccaaaat  | atcataacaa | 840  |
| ctccgcccca  | ttgacgcaaa | tagacagtag | gcatatacaa | taggaggtct  | atataaggag | 900  |
| agctcgttta  | gtgaaccgtc | agategeetg | gagacgccat | ccacactatt  | ttgacctcca | 960  |
| tagaagacac  | cgggaccgat | ccaqcctccq | caaccaaaaa | coatgoatta  | gaacccca   | 1020 |
| tccccqtqcc  | aagagtgacg | taagtaccgc | ctatagactc | tatagggacag | cocctttaaa | 1020 |
| tcttatgcat  | gctatactgt | ttttaactta | gggcctatac | acccccatt   | cettatesta | 1140 |
| taggtgatgg  | tatagettag | cctataggtg | tagattatta | accettatta  | accategora | 1200 |
| tattqqtqac  | gatactttcc | attactaatc | cataacatoo | ctctttacca  | accaccecee | 1200 |
| tattggctat  | atgccaatac | totatoctto | agagactgac | acquactetq  | tattttta   | 1200 |
| qqatqqqqtc  | ccatttatta | tttacaaatt | cacatataca | acggactetg  | cacccctaca | 1320 |
| cgcagtttt   | attaaacata | acatagaata | tccacacaca | teteggetag  | gtattagge  | 1380 |
| cataggetet  | tctccggtag | caacaasact | tecaegegaa | accetacta   | gegeeeegga | 1440 |
| agcggctcat  | ggtcgctcgg | cageteette | ctcctaacag | tagaaaaaa   | ccatgeetee | 1500 |
| agcacaatgc  | ggeegeegg  | cagteteeeg | Gagaaggag  | tggaggccag  | acctaggcac | 1560 |
| gaaaatgage  | ccaccaccac | gastsassa  | cacaaggeeg | cggcggcagg  | grargrarer | 1620 |
| deadagage   | gtggagattg | ggcccgcacg | geegaegeag | acggaagact  | taaggcagcg | 1680 |
| attacaatac  | atgcaggcag | ggagggggt  | gractetgat | aagagccaga  | ggtaactccc | 1740 |
| cacacascas  | tgttaacggt | gyagygcagt | gragrergag | cagtactcgt  | tgctgccgcg | 1800 |
| tacaataaca  | gacataatag | ccgacagact | aacagactgt | tccctccat   | gggtcttttc | 1860 |
| gcaatttccc  | gtcgtcgacg | ataattaa   | gaggeettgg | tccatgggga  | aagaagcccc | 1920 |
| canctataga  | catggctcaa | grgeateagg | ggetgatgee | aactgctccc  | cgagaggacc | 1980 |
| gagaaagga   | tctgctaaag | haceacacge | agttgggcaa | gcagcagaga  | gaaaagcaga | 2040 |
| tatttaasaa  | agagaagcct | tacaaggagg | tgacagagga | tttgctgcac  | ctcaattctc | 2100 |
| accettggagg | agaccagtag | cactgcate  | gatcgagggc | cagcccgtgg  | aggtgcttct | 2160 |
| agacaccgge  | gccgacgaca | geacegegae | cggcatcgag | cragacccac  | actacacccc | 2220 |
| gaagategtg  | ggcggcatcg | geggetteat | caacaccgga | gagtacaaga  | acgtggagat | 2280 |
| changere    | ggcaagcgca | tcaagggcac | catcatgacc | ggcgacacgc  | ccatcaacat | 2340 |
| ccceggeege  | aacctgctga | ccgccctggg | catgagcctg | aacttcccca  | tcgccaaggt | 2400 |
| ggageeegtg  | aaggtggccc | tgaageeegg | caaggacggc | cccaagetga  | agcagtggcc | 2460 |
| ceegageaag  | gagaagatcg | tggccctgcg | cgagatctgc | gagaagatgg  | agaaggacgg | 2520 |
| ccagetggag  | gaggccccac | ccaccaaccc | ctacaacacg | cccaccttcg  | ccatcaagaa | 2580 |
| gaayyacaay  | aacaagtggc | gcatgctgat | cgacttccgc | gagctgaacc  | gcgtgaccca | 2640 |
| ggactteace  | gaggtgcagc | tgggcatccc | gcaccccgcc | ggcctggcca  | agcgcaagcg | 2700 |
| catcaccgtg  | ctggacatcg | gcgacgccta | cttcagcatc | cccctggacg  | aggagttccg | 2760 |
| ccagtacacc  | gccttcaccc | tgcccagcgt | gaacaacgcc | gagcccggca  | agcgctacat | 2820 |
| ctacaaggtg  | ctgccccagg | gctggaaggg | cagccccgcc | atcttccagt  | acaccatgcg | 2880 |
| ccacgcgctg  | gagcccttcc | gcaaggccaa | ccccgacgtg | accctggtgc  | agtacatgca | 2940 |
| tgacatcetg  | atcgccagcg | accgcaccga | cctggagcac | gaccgcgtgg  | tgctgcagag | 3000 |
| caaggagetg  | ctgaacagca | teggetteag | cacgcccgag | gagaagttcc  | agaaggaccc | 3060 |
| accettecag  | tggatgggct | acgagetgtg | gcccaccaag | tggaagctgc  | agaagatcga | 3120 |
| gctgccccag  | cgcgagacct | ggaccgtgaa | cgacatccag | aagctggtgg  | gcgtgctgaa | 3180 |
| ctgggccgcc  | cagatctacc | ccggcatcaa | gaccaagcac | ctgtgccgcc  | tgatccgcgg | 3240 |
| caagatgacc  | ctgaccgagg | aggtgcagtg | gaccgagatg | gccgaggccg  | agtacgagga | 3300 |
| gaacaagatc  | atcctgagcc | aggagcagga | gggctgctac | taccaggagg  | gcaagcccct | 3360 |
| ggaggccacc  | gtgatcaaga | gccaggacaa | ccagtggagc | tacaagatcc  | accaggagga | 3420 |
| caagateetg  | aaggtgggca | agttcgccaa | gatcaagaac | acccacacca  | acggcgtgcg | 3480 |
| cctgctggcc  | cacgtgatcc | agaagatcgg | caaggaggcc | atcgtgatct  | ggggccaggt | 3540 |
| gcccaagttc  | cacctgcccg | tggagaagga | cgtgtgggag | cagtggtgga  | ccgactactg | 3600 |
| gcaggtgacc  | tggatacccg | agtgggactt | catcagcacc | cctcccctgg  | tgcgcctggt | 3660 |
| gttcaacctg  | gtgaaggacc | ccatcgaggg | cgaggagacc | tactacaccg  | acggcagctg | 3720 |
| caacaagcag  | agcaaggagg | gcaaggccgg | ctacatcacc | gaccgcggca  | aggacaaggt | 3780 |
| gaaggtgctg  | gagcagacca | ccaaccagca | ggccgagctc | gaggccttcc  | tgatggccct | 3840 |
| gaccgacagc  | ggccccaagg | ccaacatcat | cgtggacagc | cagtacgtga  | tgggcatcat | 3900 |
| caccggctgc  | cccaccgaga | gcgagagccg | cctggtgaac | cagatcatcg  | aggagatgat | 3960 |
| caagaagagc  | gagatctacg | tggcctgggt | gcccgcccac | aagggcatcg  | gcggcaacca | 4020 |

|            |            |            |            | •          |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | agaagatcga |      |
|            |            |            |            |            | tgttcaagtt |      |
|            |            |            |            |            | gccaccagaa |      |
|            |            |            |            |            | tggcctgcac |      |
|            |            |            |            |            | tcatcgaggc |      |
|            |            |            |            |            | agctggccgg |      |
|            |            |            |            |            | gccaggaggt |      |
|            |            |            |            |            | acaacccgca |      |
|            |            |            |            |            | accgcatccg |      |
|            |            |            |            |            | tgaacttcaa |      |
| gcgccgcggc | ggcatcggcg | acatgacgcc | cgccgagcgc | ctgatcaaca | tgatcaccac | 4680 |
|            |            |            |            |            | tccgcgtgta |      |
|            |            |            |            |            | ggaagggcga |      |
|            |            |            |            |            | gcaaggccaa |      |
|            |            |            |            |            | tggaggacac |      |
|            |            |            |            |            | cggatccaga |      |
| tctgctgtgc | cttctagttg | ccagccatct | gttgtttgcc | cctcccccgt | gccttccttg | 5040 |
|            |            |            |            |            | tgcatcgcat |      |
|            |            |            |            |            | caagggggag |      |
| gattgggaag | acaatagcag | gcatgctggg | gatgcggtgg | gctctatggg | tacccaggtg | 5220 |
| ctgaagaatt | gacccggttc | ctcctgggcc | agaaagaagc | aggcacatcc | ccttctctgt | 5280 |
|            |            |            |            |            | acactcatag |      |
| ctcaggaggg | ctccgccttc | aatcccaccc | gctaaagtac | ttggagcggt | ctctccctcc | 5400 |
| ctcatcagcc | caccaaacca | aacctagcct | ccaagagtgg | gaagaaatta | aagcaagata | 5460 |
| ggctattaag | tgcagaggga | gagaaaatgc | ctccaacatg | tgaggaagta | atgagagaaa | 5520 |
| tcatagaatt | ttaaggccat | catggcctta | atcttccgct | tectegetea | ctgactcgct | 5580 |
| gcgctcggtc | gttcggctgc | ggcgagcggt | atcagctcac | tcaaaggcgg | taatacggtt | 5640 |
| atccacagaa | tcaggggata | acgcaggaaa | gaacatgtga | gcaaaaggcc | agcaaaaggc | 5700 |
|            |            |            |            |            | cccctgacga |      |
|            |            |            |            |            | tataaagata |      |
|            |            |            |            |            | tgccgcttac |      |
| cggatacctg | tccgcctttc | tecetteggg | aagcgtggcg | ctttctcata | gctcacgctg | 5940 |
| taggtatctc | agttcggtgt | aggtcgttcg | ctccaagctg | ggctgtgtgc | acgaaccccc | 6000 |
|            |            |            |            |            | acccggtaag |      |
| acacgactta | tcgccactgg | cagcagccac | tggtaacagg | attagcagag | cgaggtatgt | 6120 |
|            |            |            |            |            | gaagaacagt |      |
| atttggtatc | tgcgctctgc | tgaagccagt | taccttcgga | aaaagagttg | gtagctcttg | 6240 |
|            |            |            |            |            | agcagattac |      |
|            |            |            |            |            | ctgacgctca |      |
| gtggaacgaa | aactcacgtt | aagggatttt | ggtcatgaga | ttatcaaaaa | ggatcttcac | 6420 |
| ctagatcctt | ttaaattaaa | aatgaagttt | taaatcaatc | taaagtatat | atgagtaaac | 6480 |
|            |            |            |            |            | tctgtctatt |      |
| tcgttcatcc | atagttgcct | gactcggggg | gggggggcgc | tgaggtctgc | ctcgtgaaga | 6600 |
|            |            |            |            |            | aagtgaggga |      |
|            |            |            |            |            | acttttgctt |      |
| tgccacggaa | cggtctgcgt | tgtcgggaag | atgcgtgatc | tgatccttca | actcagcaaa | 6780 |
| agttcgattt | attcaacaaa | gccgccgtcc | cgtcaagtca | gcgtaatgct | ctgccagtgt | 6840 |
|            |            |            |            |            | aaactgcaat |      |
|            |            |            |            |            | taatgaagga |      |
|            |            |            |            |            | tgcgattccg |      |
| actcgtccaa | catcaataca | acctattaat | ttcccctcgt | caaaaataag | gttatcaagt | 7080 |
| gagaaatcac | catgagtgac | gactgaatcc | ggtgagaatg | gcaaaagctt | atgcatttct | 7140 |
| ttccagactt | gttcaacagg | ccagccatta | cgctcgtcat | caaaatcact | cgcatcaacc | 7200 |
|            |            |            |            |            | gctgttaaaa |      |
| ggacaattac | aaacaggaat | cgaatgcaac | cggcgcagga | acactgccag | cgcatcaaca | 7320 |
|            |            |            |            |            |            |      |

```
atatttteac ctgaateagg atattettet aatacetgga atgetgttt cceggggate 7380 geagtggtga gtaaceatge ateateagga gtaeggataa aatgettgat ggteggaaga 7440 ggcataaatt cegteageca gtttagtetg aceateteat etgtaacate aateggatag 7500 etacetttge catgtteag aaacaactet ggegeategg getteecata caateggatag 7560 attgtegeac etgattgeec gacattateg eggeecatt tatacecata taaateagea 7620 teeatgttgg aatttaateg eggeetegag caagaegttt ecegttgaat atggeteata 7680 acacecettg tattactgtt tatgtaagea gacagtttta ttgtteatga tgatatattt 7740 ttatettgtg eaatgtaaca teagagattt tgagaecaca egtggettte eceeceeee 7800 eattattgaa geatttatea gggttattgt eteatgageg gatacatatt tgaatgtatt 7860 tagaaaaata aacaaatagg ggtteegege acattteee gaaaagtgee acettgaege gaggeettte eggetateae gggttatege eacatteeee gaaaagtgee aecetgaege 7920 taagaaaaca ttattateat gacattaace tataaaaaata ggegtateae gaggeettte 7980 egte
```

<210> 163
<211> 9194
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 163

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtct gtaageggat geegggagca gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattqta ctqaqaqtqc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatq 300 tocaacatta cogocatgit gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacqtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataaqcaq 900 agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac egggacegat ecageeteeg eggeegggaa eggtgeattg gaaegeggat 1020 teccegtgee aagagtgaeg taagtacege etatagaete tataggeaea eecetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet beteeggtag eggeggaget becacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg eageteettg eteetaaeag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860

|            |            |            |            | ctgagcggcg |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            | aagtacaagc |            |      |
|            |            |            |            | cccggcctgc |            |      |
| cgagggctgc | cgccagatcc | tgggccagct | gcagcccagc | ctgcagaccg | gcagcgagga | 2100 |
| gctgcgcagc | ctgtacaaca | ccgtggccac | cctgtactgc | gtgcaccagc | gcatcgagat | 2160 |
| caaggacacc | aaggaggccc | tggacaagat | cgaggaggag | cagaacaaga | gcaagaagaa | 2220 |
| ggcccagcag | gccgccgccg | acaccggcca | cagcaaccag | gtgagccaga | actaccccat | 2280 |
| cgtgcagaac | atccagggcc | agatggtgca | ccaggccatc | agcccccgca | ccctgaacgc | 2340 |
| ctgggtgaag | gtggtggagg | agaaggcctt | cagccccgag | gtgatcccca | tgttcagcgc | 2400 |
| cctgagcgag | ggcgccaccc | cccaggacct | gaacaccatg | ctgaacaccg | tgggcggcca | 2460 |
| ccaggccgcc | atgcagatgc | tgaaggagac | catcaacgag | gaggccgccg | agtgggaccg | 2520 |
| cgtgcacccc | gtgcacgccg | gccccatcgc | ccccggccag | atgcgcgagc | cccgcggcag | 2580 |
| cgacatcgcc | ggcaccacca | gcaccetgca | ggagcagatc | ggctggatga | ccaacaaccc | 2640 |
| ccccatcccc | gtgggcgaga | tctacaagcg | ctggatcatc | ctgggcctga | acaagatcgt | 2700 |
| gcgcatgtac | agccccacca | gcatcctgga | catccgccag | ggccccaagg | agcccttccg | 2760 |
| cgactacgtg | gaccgcttct | acaagaccct | gcgcgccgag | caggccagcc | aggaggtgaa | 2820 |
| gaactggatg | accgagaccc | tgctggtgca | gaacgccaac | cccgactgca | agaccatcct | 2880 |
| gaaggccctg | ggccccgccg | ccaccctgga | ggagatgatg | accgcctgcc | agggcgtggg | 2940 |
| cggccccggc | cacaaggccc | gcgtgctggc | cgaggccatg | agccaggtga | ccaacagcgc | 3000 |
|            |            |            |            | aagatcgtga |            |      |
|            |            |            |            | ccccgcaaga |            |      |
|            |            |            |            | gagcgacagg |            |      |
|            |            |            |            | tttcttcaga |            |      |
|            |            |            |            | gagacaacaa |            |      |
|            |            |            |            | teceteagat |            |      |
|            |            |            |            | aggaggccct |            |      |
|            |            |            |            | gccgctggaa |            |      |
|            |            |            |            | accagatect |            |      |
|            |            |            |            | ccccgtgaa  |            |      |
|            |            |            |            | ccatcagccc |            |      |
|            |            |            |            | tgaagcagtg |            |      |
|            |            |            |            | tggagaagga |            |      |
|            |            |            |            | tcgccatcaa |            |      |
|            |            |            |            | acaagcgcac |            |      |
|            |            |            |            | agcagaagaa |            |      |
|            |            |            |            | acaaggactt |            |      |
|            |            |            |            | gcatccgcta |            |      |
|            |            |            |            | agtgcagcat |            |      |
|            |            |            |            | accagtacat |            |      |
|            |            |            |            | agatcgagga |            |      |
|            |            |            |            | accagaagga |            |      |
|            |            |            |            | tgcagcccat |            |      |
|            |            |            |            | tgggcaagct |            |      |
|            |            |            |            | agctgctgcg |            |      |
|            |            |            |            | tggagctggc |            |      |
|            |            |            |            | ccagcaagga |            |      |
|            |            |            |            | tctaccagga |            |      |
|            |            |            |            | acaccaacga |            |      |
|            |            |            |            | tgatctgggg |            |      |
|            |            |            |            | ggtggaccga |            |      |
|            |            |            |            | ccctggtgaa |            |      |
|            |            |            |            | acgtggacgg |            |      |
|            |            |            |            | geggeegeea |            |      |
|            |            |            |            | ccatccacct |            |      |
|            |            |            |            | acgccctggg |            |      |
| J          |            |            | J          | - 5 - 555  |            |      |

|            |            |            |            |            | gctgatcaag |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | caacgagcag |      |
|            |            |            |            |            | catcgacaag |      |
|            |            |            |            |            | cgacttcaac |      |
|            |            |            |            |            | gctgaagggc |      |
| gaggccatgc | acggccaggt | ggactgcagc | cccggcatct | ggcagctgga | ctgcacccac | 5520 |
| ctggagggca | aggtgatcct | ggtggccgtg | cacgtggcca | gcggctacat | cgaggccgag | 5580 |
| gtgatccccg | ccgagaccgg | ccaggagacc | gcctacttcc | tgctgaagct | ggccggccgc | 5640 |
| tggcccgtga | agaccgtgca | caccgacaac | ggcagcaact | tcaccagcac | caccgtgaag | 5700 |
|            |            |            |            |            | ccccagagc  |      |
|            |            |            |            |            | ggtgcgcgac |      |
|            |            |            |            |            | cttcaagcgc |      |
|            |            |            |            |            | cgccaccgac |      |
|            |            |            |            |            | cgtgtactac |      |
|            |            |            |            |            | gggcgagggc |      |
|            |            |            |            |            | ggccaagatc |      |
|            |            |            |            |            | ccaggacgag |      |
|            |            |            |            |            | cctccccgt  |      |
|            |            |            |            |            | atgaggaaat |      |
|            |            |            |            |            | ggcaggacag |      |
|            |            |            |            |            | gctctatggg |      |
|            |            |            |            |            | aggcacatcc |      |
|            |            |            |            |            | cactcatagg |      |
|            |            |            |            |            | ttggagcggt |      |
|            |            |            |            |            | gaagaaatta |      |
|            |            |            |            |            | tgaggaagta |      |
|            |            |            |            |            | tcctcgctca |      |
|            |            |            |            |            | tcaaaggcgg |      |
|            |            |            |            |            | gcaaaaggcc |      |
|            |            |            |            |            | aggctccgcc |      |
|            |            |            |            |            | ccgacaggac |      |
|            |            |            |            |            | gttccgaccc |      |
|            |            |            |            |            | ctttctcata |      |
|            |            |            |            |            | ggctgtgtgc |      |
|            |            |            |            |            | cttgagtcca |      |
|            |            |            |            |            | attagcagag |      |
|            |            |            |            |            | ggctacacta |      |
|            |            |            |            |            | aaaagagttg |      |
|            |            |            |            |            | gtttgcaagc |      |
|            |            |            |            |            | tctacggggt |      |
|            |            |            |            |            | ttatcaaaaa |      |
|            |            |            |            |            | taaagtatat |      |
|            |            |            |            |            | atctcagcga |      |
|            |            |            |            |            | tgaggtctgc |      |
|            |            |            |            |            | tccagccaga |      |
|            |            |            |            |            | gtgattttga |      |
|            |            |            |            |            | tgatccttca |      |
|            |            |            |            |            | gcgtaatgct |      |
|            |            |            |            |            | gcatcaaatg |      |
|            |            |            |            |            | gccgtttctg |      |
|            |            |            |            |            | ggtatcggtc |      |
|            |            |            |            |            | caaaaataag |      |
|            |            |            |            |            | gcaaaagctt |      |
|            |            |            |            |            | caaaatcact |      |
|            |            |            |            |            | atacgcgatc |      |
| -          |            | 2.2        |            |            |            |      |

```
getgttaaaa ggacaattac aaacaggaat egaatgcaac eggegeagga acactgceag 8520 egcatcaaca atattteac etgaatcagg atattette aatacetgga atgetgttt 8580 eeggggate geagtggtga gtaaccatge atcateagga gtaeggataa aatgettgat 8640 ggteggaaga ggcataaatt eegteageca gtttagtetg aceateteat etgtaacate 8700 attggeaaeg etaecettge eatgtteag aaacaactet ggegeategg getteecata 8760 eaategatag attgtegeae etgattgee gacattateg eggeetetga eagacett tataeceata 8820 taaatcagea tecatgttgg aatttaateg eggeetegag eaagaegttt eeegttgaat 8880 atggeteata acaceeettg tattaetgtt tatgtaagea gacagtttta ttgtteatga 8940 tgatatattt ttatettgtg eaatgtaaca teagagattt tgagacacaa egtggettee 9000 eeeeeeeeee eattattgaa geatttatea gggttattgt eteatgaeg gataecatatt 9060 tgaatgtatt taagaaaata aacaaatagg ggtteegege acattteeee gaaaagtgee 9120 acetgaegte egte
```

<210> 164
<211> 9194
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 164

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cocqcctqqc tgaccqccca acqacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getataetgt tittggetig gggeetatae acceegett cettatgeta 1140 taggtgatgg tatagcttag cetataggtg tgggttattg accattattg accactecee 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 qcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgeggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800

|              |              |             |            |                      |                          | 1060 |
|--------------|--------------|-------------|------------|----------------------|--------------------------|------|
| cgcgccacca   | gacataatag   | ctgacagact  | aacagactgt | teettteeat           | gggcccccc                | 1000 |
| tgcagtcacc   | grcgrcgaca   | ragacacca   | cgccagcgcg | ctgageggeg           | gegagergga               | 1920 |
| ccgctgggag   | aagatccgcc   | racacccaa   | cggcaagaag | aagtacaage           | tgaagcacat               | 1980 |
| cgtgtgggcc   |              |             |            |                      |                          |      |
| cgagggctgc   | cgccagatcc   | tgggccagct  | gcagcccagc | ctgcagaccg           | gcagcgagga               | 2100 |
| gctgcgcagc   | ctgtacaaca   | ccgtggccac  | cctgtactgc | gtgcaccagc           | gcatcgagat               | 2160 |
| caaggacacc   | aaggaggccc   | tggacaagat  | cgaggaggag | cagaacaaga           | gcaagaagaa               | 2220 |
| ggcccagcag   | geegeegeeg   | acaccggcca  | cagcaaccag | gtgagccaga           | actaccccat               | 2280 |
| cgtgcagaac   | atccagggcc   | agatggtgca  | ccaggccatc | agcccccgca           | ccctgaacgc               | 2340 |
| ctgggtgaag   | gtggtggagg   | agaaggcctt  | cagccccgag | gtgatcccca           | tgttcagcgc               | 2400 |
| cctgagcgag   | ggcgccaccc   | cccaggacct  | gaacaccatg | ctgaacaccg           | tgggcggcca               | 2460 |
| ccaggccgcc   | atgcagatgc   | tgaaggagac  | catcaacgag | gaggccgccg           | agtgggaccg               | 2520 |
| cgtgcacccc   | gtgcacgccg   | gccccatcgc  | ccccggccag | atgcgcgagc           | cccgcggcag               | 2580 |
| cgacatcgcc   | ggcaccacca   | gcaccctgca  | ggagcagatc | ggctggatga           | ccaacaaccc               | 2640 |
| ccccatcccc   | qtqqqqaga    | tctacaagcg  | ctggatcatc | ctgggcctga           | acaagatcgt               | 2700 |
| gcgcatgtac   | aqccccacca   | gcatcctgga  | catccgccag | ggccccaagg           | agcccttccg               | 2760 |
| caactacata   | gaccgcttct   | acaagaccct  | gcgcgccgag | caggccagcc           | aggaggtgaa               | 2820 |
| gaactggatg   | accgagaccc   | tactaataca  | gaacgccaac | cccgactgca           | agaccatcct               | 2880 |
| gaaggcctg    | aaccccacca   | ccaccctqqa  | ggagatgatg | accocctocc           | agggcgtggg               | 2940 |
| caacccaac    | cacaaggccc   | acatactaac  | cgaggccatg | agccaggtga           | ccaacagcgc               | 3000 |
| caccatcato   | atocaococo   | gcaacttccq  | caaccagcgc | aaqatcqtqa           | agtgcttcaa               | 3060 |
| chacaacaaa   | gagggccaca   | ccacccacaa  | ctaccacacc | ccccacaaqa           | agggctgctg               | 3120 |
| daautucaac   | aaggagggc    | accagatgaa  | ggactgcacc | gagegaeagg           | ctaattttt                | 3180 |
| addaadatc    | tagcetteee   | acaagggaag  | gccagggaat | tttcttcaga           | gcagaccaga               | 3240 |
| agggaagacc   | caccadaad    | agagetteag  | atttaagaaa | gagacaacaa           | ctccctctca               | 3300 |
| gecaacagee   | ccaetagaag   | aggaactgta  | teetttaget | teceteagat           | cactctttgg               | 3360 |
| gaagcaggag   | tegteagaea   | aaadataggg  | aaccaactaa | aggaggggt            | tctagacacc               | 3420 |
| cagogacccc   | acaccatact   | qqaqqaqatq  | aacctgcccg | accactagaa           | gcccaagatg               | 3480 |
| ggcgccgacg   | teggegett    | catcaacctc  | aaccaataca | accadatect           | gatcgagatc               | 3540 |
| tacaaccaca   | acaccatca    | caccatacta  | ataaacccca | ccccatasa            | catcatcggc               | 3600 |
| cgcggccaca   | tgaccacaga   | caachacacc  | ctgaacttcc | ccatcagccc           | catcgagacc               | 3660 |
| atagagatas   | agetgaagee   | cancatage   | aacccaaaa  | tgaagcagtg           | gcccctgacc               | 3720 |
| gegeeegea    | tcaagggggt   | agtagagate  | tacaccaaga | tagagaagaa           | gggcaagatc               | 3780 |
| gaggagaaga   | accccasass   | ccctacaac   | accccatat  | togocatosa           | gaagaaggac               | 3840 |
| agcaagaccg   | gccccgagaa   | agtagacttc  | cacaaactaa | acaagcgcac           | ccaggacttc               | 3900 |
| tagaacataa   | agetgggeat   | ccccaccc    | accaacctaa | agcagaagaa           | gagcgtgacc               | 3960 |
| cgggaggcgc   | tagacaecae   | ctacttcacc  | ataccctaa  | acaaggactt           | ccgcaagtac               | 4020 |
| gegeegbeeg   | cgggcgacgc   | catcaacaac  | Pedececea  | gcatcocta            | ccagtacaac               | 4080 |
| attactacca   | aggachagaa   | adacadecee  | gccatcttcc | agtgcagcat           | gaccaagatc               | 4140 |
| gtgttgttt    | tegggeeggaa  | gageageece  | atcotoatct | accadtacat           | ggaccacctg               | 4200 |
| · tacgtgggca | accacctage   | gaacecegae  | caccacacca | agatrgagga           | actacaccaa               | 4260 |
| cacytygyca   | gcgacccgga   | gaccagccag  | caccacaca  | agacogagga           | gcccccttc                | 4320 |
| caccigotge   | gerggggeet   | acacccccc   | aactagaage | tacaacccat           | cgtgctgccc               | 4380 |
| ctgtggatgg   | getacgaget   | geaccecgae  | cagaagetgg | tagacaaact           | gaactgggcc               | 4440 |
| gagaaggaca   | acceptacege  | caacgacacc  | cagaageegg | arctactaca           | cggcaccaag               | 4500 |
| agecagacec   | acgeeggeat   | cataggegege | gaggggggg  | tagaactaac           | cgagaaccgc               | 4560 |
| geeetgaeeg   | aggragaget   | aceasacata  | tactaccacc | caacaacaa            | cctgatcgc                | 4620 |
| gagaccccga   | aggageeege   | addecaates  | acchaccaca | tctaccacca           | cctgatcgcc<br>gcccttcaag | 4680 |
| gagacccaga   | ageagggeea   | 222ccaaraa  | accuactaga | acaccayya            | catassocse               | 4740 |
| aacccgaaga   | ceggeaagea   | gategeatg   | aagggegeee | tastataaa            | cgtgaagcag               | 4800 |
| ergacegagg   | tegegeagaa   | gattgctatt  | tagasaast  | catacacca            | caagaccccc               | 4860 |
| aaguccaage   | tgoccateca   | gaaggagacc  | 2333433666 | ccctcctca            | gtactggcag<br>gctgtggtac | 4920 |
| geeacetgga   | - ceecegageg | categgegeg  | gagagettet | acataascaa           | cgccgccaac               | 4980 |
| cagerggaga   | aggageeeat   | daccaactec  | gagacccccc | acaccacca<br>acacaga | gaaggtggtg               | 5040 |
| egegagaeea   | agetgggeaa   | agecagecae  | gagactgact | ccatccacct           | ggccctgcag               | 5100 |
| eccetgaceg   | acaccaccaa   | Clagaagact  | gagergeagg | Juliundanu           | aaaaaaaaa                | 2200 |

|            | taasaatass   | catcotoacc   | gacagccagt   | acgccctggg   | catcatccag                               | 5160 |
|------------|--------------|--------------|--------------|--------------|--|------|
| gacageggee | rggaggrgaa   | nagrgagetg   | ataaaccaga   | tcatcgagca   | gctgatcaag                               | 5220 |
| geeeageeeg | tatacataac   | chagataccc   | gcccacaagg   | gcatcggcgg   | caacqagcag                               | 5280 |
| aaggagaagg | tactaccage   | cadeatecae   | aaggtgctgt   | tcctggacgg   | catcgacaag                               | 5340 |
| geggaeggee | ragraagege   | ctaccacac    | aactggcgcg   | ccatggccag   | cgacttcaac                               | 5400 |
| gcccaggagg | agcacyagaa   | gcaccacage   | accepted     | acaagtgcca   | actaaaagac                               | 5460 |
| ctgcccccg  | tggtggccaa   | ggagaccgcg   | cccagcatct   | ggcagctggc   | atocacccac                               | 5520 |
| gaggccatgc | acggccaggt   | ggaccgcage   | cacataacca   | ggcagctacat  | cgaggccgag                               | 5580 |
| ctggagggca | aggtgatect   | ggtggttgt    | acctacttcc   | gcggctacat   | aaccaaccac                               | 5640 |
| gtgatccccg | cegagacegg   | ccaggagacc   | gacagaact    | tgctgaagct   | caccataaaa                               | 5700 |
| tggcccgtga | agaccgtgca   | caccyacaac   | ggcagcaacc   | tcaccagcac   | ccccagage                                | 5760 |
| accacctact | ggtgggccgg   | cattaagtag   | gageeeggea   | tcccctacaa   | aatacacaac                               | 5820 |
| cagggcgtga | ccgagageac   | gaacaaggag   | ataccatat    | tcatcggcca   | cttcaagcgc                               | 5880 |
| caggccgagc | acctgaagac   | egeegegeag   | arggeegeste  | tcatccacaa   | caccaccasc                               | 5940 |
| aagggcggca | teggeggeta   | cagegeegge   | gagegeateg   | tggacatcat   | catatactac                               | 6000 |
| atccagacca | aggagetgea   | gaagcagacc   | accaagaccc   | agaacttccg   | adacasadac                               | 6060 |
| cgcgacagcc | gcgaccccgt   | geggaaggge   | cccgccaage   | tgctgtggaa   | adccasastc                               | 6120 |
| gccgtggtga | tccaggacaa   | cagcgacatc   | aaggugguge   | cccgccgcaa   | ggccaagacc                               | 6180 |
| atccgcgact | acggcaagca   | gatggccggc   | gacgactgcg   | tggccagccg   | cetececet                                | 6240 |
| gactaggaat | tetgetgtge   | cttctagttg   | ccagccatct   | gttgtttgcc   | atazaazat                                | 6300 |
| gccttccttg | accctggaag   | gtgccactcc   | cactgtcctt   | tcctaataaa   | acgaggaaac                               | 6360 |
| tgcatcgcat | tgtctgagta   | ggtgtcattc   | tattctgggg   | ggtggggtgg   | ggcaggacag                               | 6420 |
| caagggggag | gattgggaag   | acaatagcag   | geatgetggg   | gatgcggtgg   | aggagata                                 | 6480 |
| tacccaggtg | ctgaagaatt   | gacccggttc   | ctcctgggcc   | agaaagaagc   | aggeacacce                               | 6540 |
| ccttctctgt | gacacaccct   | gtccacgccc   | ctggttctta   | gttccagccc   | btecacage                                | 6600 |
| acactcatag | ctcaggaggg   | ctccgccttc   | aatcccaccc   | gctaaagtac   | ccggagcggc                               | 6660 |
| ctctccctcc | ctcatcagcc   | caccaaacca   | aacctagcct   | ccaagagtgg   | bassassassassassassassassassassassassass | 6720 |
| aagcaagata | ggctattaag   | tgcagaggga   | gagaaaatgc   | ctccaacatg   | tgaggaagca                               | 6720 |
| atgagagaaa | tcatagaatt   | ttaaggccat   | catggcctta   | atcttccgct   | teetegetea                               | 6940 |
| ctgactcgct | gcgctcggtc   | gttcggctgc   | ggcgagcggt   | atcagctcac   | ccaaaggcgg                               | 6000 |
| taatacggtt | atccacagaa   | tcaggggata   | acgcaggaaa   | gaacatgtga   | gcaaaaygcc                               | 6960 |
| agcaaaaggc | caggaaccgt   | aaaaaggccg   | , cgttgctggc | gtttttccat   | aggetteeget                              | 7020 |
| cccctgacga | gcatcacaaa   | aatcgacgct   | caagtcagag   | gtggcgaaac   | ccgacaggac                               | 7020 |
| tataaagata | ccaggcgttt   | cccctggaa    | geteectege   | gegetetect   | gereegaeee                               | 7140 |
| tgccgcttac | cggatacctg   | teegeettt    | tecetteggg   | aagcgtggcg   | ecetetetee                               | 7200 |
| gctcacgctg | taggtatctc   | agttcggtgt   | , aggtcgttcg | ctccaagctg   | ggergrates                               | 7260 |
| acgaaccccc | cgttcagccc   | gaccgctgcg   | g ccttatccgg | taactatcgt   | ettgagteta                               | 7220 |
| acccggtaag | acacgactta   | tegecactge   | g cagcagccac | tggtaacagg   | actageagag                               | 7320 |
| cgaggtatgt | . aggcggtgct | : acagagttc  | tgaagtggtg   | gcctaactac   | ggctacacta                               | 7440 |
| gaagaacagt | . atttggtato | : tgcgctctg  | tgaagccagt   | taccttcgga   | aaaagagttg                               | 7500 |
| ghagetette | , atccqqcaaa | a caaaccacco | g ctggtagcgg | , rggrrrrrr  | gilligeaage                              | 7300 |
| aggagattag | · ococadaaaa | , aaaggatete | aagaagatco   | : tttgatctt  | cctacggggt                               | 1500 |
| chasagetes | ı ataqaacqaa | a aactcacqti | t aagggatttt | : ggtcatgaga | CLaccadada                               | 7020 |
| ggatetteac | : ctagatcctt | : ttaaattaa: | a aatgaagtti | taaatcaatc   | taaagtatat                               | 7000 |
| atgagtaaad | : ttggtctgad | agttaccaa    | t gottaatcag | g tgaggcacct | acctcagcga                               | 7/40 |
| tctgtctatt | tcgttcatco   | c atagttgcc  | t gactcggggg | g gggggggg   | tgaggtetge                               | 7960 |
| ctcataaaga | a addtattaci | t qactcatac  | c aggcctgaat | c cgccccatca | cccagccaga                               | 7000 |
| aagtgaggg  | a accacaatt  | r atgagaget  | t tgttgtagg! | t ggaccagttg | gegaelelga                               | 1920 |
| actittacti | toccacqqa:   | a cogtetoco  | t tgtcgggaag | g atgcgtgatc | tgatecttea                               | 1900 |
| actcagcaaa | a agttcgatt  | t attcaacaa  | a geegeegted | c cgtcaagtca | gegtaatget                               | 8040 |
| chaccagia  | t tacaaccaa  | t taaccaatt  | c tgattagaaa | a aactcatcga | gcatcaaatg                               | BTOO |
| ssactocaat | t ttattcata  | t caggattat  | c aataccatai | t ttttgaaaaa | geegtttetg                               | 8700 |
| taatgaagg  | a daaaactca  | c caaqqcaqt  | t ccataggate | g gcaagatcct | : ggtatcggtc                             | 8220 |
| tacaattcc  | acteqteca:   | a catcaatac  | a acctattaa  | t ttcccctcgt | : caaaaacaag                             | 0200 |
| ottatcaag  | t gagaaatca  | c catqaqtqa  | c gactgaatc  | c ggtgagaatg | g gcaaaagctt                             | 8340 |
| atgcatttc  | t ttccagact  | t gttcaacag  | g ccagccatt: | a cgctcgtcat | caaaatcact                               | 8400 |
|            |              |              |              |              |  |      |

PCT/US01/25721 WO 02/32943

```
cgcatcaacc aaaccqttat tcattcgtqa ttgcgcctga qcgagacgaa atacgcgatc 8460
gctgttaaaa ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag 8520
cgcatcaaca atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt 8580
cccggggatc gcagtggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat 8640
ggtcggaaga ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc 8700
attqqcaacq ctacctttgc catgtttcaq aaacaactct ggcgcatcgg gcttcccata 8760
caatcgatag attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata 8820
taaatcagca tocatqttqq aatttaatcq cqqcctcqaq caaqacgttt cccqttgaat 8880
atggctcata acaccccttg tattactgtt tatgtaagca gacagtttta ttgttcatga 8940
tgatatattt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc 9000
ccccccccc cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 9060
tqaatqtatt taqaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 9120
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 9180
gaggcccttt cqtc
                                                                  9194
```

<210> 165 <211> 12411 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300

qqqqtcatta qttcataqcc catatatqqa qttccqcqtt acataactta cqqtaaatqq 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720

<400> 165

ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 ageteqttta gtqaaccgtc agategeetg gagaegeeat ceaegetgtt ttgaccteea 960 tagaagacac cgggaccgat ccaqcctccq cggccgggaa cggtgcattg gaacgcggat 1020 tccccqtqcc aaqaqtqacg taagtaccqc ctatagactc tataggcaca cccctttggc 1080 tottatgcat gotatactgt tittggcttg gggcctatac acccccgctt cottatgcta 1140

tattggctat atgccaatac tetgteette agagactgae acggaetetg tattttaca 1320 qqatqqqqtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeg aqccetggte ceatgeetee 1500 agggggteat ggtegetegg cageteettg etectaacag tggaggecag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620

gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 qcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360

catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgbcaatggg agthtgthth ggcaccaaaa bcaacgggac thtccaaaat ghcghaacaa 840

taggtgatgq tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260

| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg | 1800 |
|------------|------------|------------|------------|------------|------------|------|
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | teettteeat | gggtctttc  | 1860 |
| tgcagtcacc | gtcgtcgaca | tgggcgcccg | cgccagcgtg | ctgagcggcg | gcgagctgga | 1920 |
| ccgctgggag | aagatccgcc | tgcgccccgg | cggcaagaag | aagtacaagc | tgaagcacat | 1980 |
| cgtgtgggcc | agccgcgagc | tggagcgctt | cgccgtgaac | cccggcctgc | tggagaccag | 2040 |
| cgagggctgc | cgccagatcc | tgggccagct | gcagcccagc | ctgcagaccg | gcagcgagga | 2100 |
| gctgcgcagc | ctgtacaaca | ccgtggccac | cctgtactgc | gtgcaccagc | gcatcgagat | 2160 |
| caaggacacc | aaggaggccc | tggacaagat | cgaggaggag | cagaacaaga | gcaagaagaa | 2220 |
| ggcccagcag | geegeegeeg | acaccggcca | cagcaaccag | gtgagccaga | actaccccat | 2280 |
| cgtgcagaac | atccagggcc | agatggtgca | ccaggccatc | agcccccgca | ccctgaacgc | 2340 |
| ctgggtgaag | gtggtggagg | agaaggcctt | cagccccgag | gtgatcccca | tgttcagcgc | 2400 |
| cctgagcgag | ggcgccaccc | cccaggacct | gaacaccatg | ctgaacaccg | tgggcggcca | 2460 |
| ccaggccgcc | atgcagatgc | tgaaggagac | catcaacgag | gaggccgccg | agtgggaccg | 2520 |
| cgtgcacccc | gtgcacgccg | gccccatcgc | ccccggccag | atgcgcgagc | cccgcggcag | 2580 |
| cgacatcgcc | ggcaccacca | gcaccctgca | ggagcagatc | ggctggatga | ccaacaaccc | 2640 |
| ccccatcccc | gtgggcgaga | tctacaagcg | ctggatcatc | ctgggcctga | acaagatcgt | 2700 |
| gcgcatgtac | agccccacca | gcatcctgga | catccgccag | ggccccaagg | agcccttccg | 2760 |
| cgactacgtg | gaccgcttct | acaagaccct | gcgcgccgag | caggccagcc | aggaggtgaa | 2820 |
| gaactggatg | accgagaccc | tgctggtgca | gaacgccaac | cccgactgca | agaccatcct | 2880 |
| gaaggccctg | ggccccgccg | ccaccctgga | ggagatgatg | accgcctgcc | agggcgtggg | 2940 |
| cggccccggc | cacaaggccc | gcgtgctggc | cgaggccatg | agccaggtga | ccaacagcgc | 3000 |
| caccatcatg | atgcagcgcg | gcaacttccg | caaccagcgc | aagatcgtga | agtgcttcaa | 3060 |
| ctgcggcaag | gagggccaca | ccgcccgcaa | ctgccgcgcc | ccccgcaaga | agggctgctg | 3120 |
| gaagtgcggc | aaggagggcc | accagatgaa | ggactgcacc | gagcgacagg | ctaattttt  | 3180 |
| agggaagatc | tggccttccc | acaagggaag | gccagggaat | tttcttcaga | gcagaccaga | 3240 |
| gccaacagcc | ccaccagaag | agagcttcag | gtttggggaa | gagacaacaa | ctccctctca | 3300 |
| gaagcaggag | ccgatagaca | aggaactgta | tcctttagct | tccctcagat | cactctttgg | 3360 |
| cagegaeece | tcgtcacaat | aaagataggg | ggccagctga | aggaggccct | tctagacacc | 3420 |
| ggcgccgacg | acaccgtgct | ggaggagatg | aacctgcccg | gccgctggaa | gcccaagatg | 3480 |
| atcggcggca | teggeggett | catcaaggtg | ggccagtacg | accagatect | gatcgagatc | 3540 |
| tgcggccaca | aggccatcgg | caccgtgctg | gtgggcccca | ccccgtgaa  | catcatcggc | 3600 |
| cgcaacctgc | tgacccagat | cggctgcacc | ctgaacttcc | ccatcagccc | catcgagacc | 3660 |
| gtgcccgtga | agctgaagcc | cggcatggac | ggccccaagg | tgaagcagtg | gcccctgacc | 3720 |
| gaggagaaga | tcaaggccct | ggtggagatc | tgcaccgaga | tggagaagga | gggcaagatc | 3780 |
| agcaagatcg | gccccgagaa | cccctacaac | acccccgtgt | tegecateaa | gaagaaggac | 3840 |
| agcaccaagt | ggcgcaagct | ggtggacttc | cgcgagctga | acaagcgcac | ccaggacttc | 3900 |
| tgggaggtgc | agctgggcat | ccccacccc  | gccggcctga | agcagaagaa | gagcgtgacc | 3960 |
| gtgctggacg | tgggcgacgc | ctacttcagc | gtgcccctgg | acaaggactt | ccgcaagtac | 4020 |
| accgccttca | ccatccccag | catcaacaac | gagacccccg | gcatccgcta | ccagtacaac | 4080 |
| gtgctgcccc | agggctggaa | gggcagcccc | gccatcttcc | agtgcagcat | gaccaagatc | 4140 |
| ctggagccct | teegeaagea | gaaccccgac | atcgtgatct | accagtacat | ggacgacctg | 4200 |
| tacgtgggca | gcgacctgga | gatcggccag | caccgcacca | agatcgagga | gctgcgccag | 4260 |
| cacctgctgc | gctggggctt | caccaccccc | gacaagaagc | accagaagga | gcccccttc  | 4320 |
| ctgtggatgg | gctacgagct | gcaccccgac | aagtggaccg | tgcagcccat | cgtgctgccc | 4380 |
| gagaaggaca | gctggaccgt | gaacgacatc | cagaagctgg | tgggcaagct | gaactgggcc | 4440 |
| agccagatct | acgccggcat | caaggtgcgc | cagctgtgca | agctgctgcg | cggcaccaag | 4500 |
| gccctgaccg | aggtggtgcc | cctgaccgag | gaggccgagc | tggagctggc | cgagaaccgc | 4560 |
| gagatcctga | aggagcccgt | gcacggcgtg | tactacgacc | ccagcaagga | cctgatcgcc | 4620 |
| gagatccaga | agcagggcca | gggccagtgg | acctaccaga | tctaccagga | gcccttcaag | 4680 |
| aacctgaaga | ccggcaagta | cgcccgcatg | aagggcgccc | acaccaacga | cgtgaagcag | 4740 |
| ctgaccgagg | ccgtgcagaa | gategecace | gagagcatcg | tgatctgggg | caagaccccc | 4800 |
| aagttcaago | tgcccatcca | gaaggagacc | tgggaggcct | ggtggaccga | gtactggcag | 4860 |
| gccacctgga | tccccgagtg | ggagttcgtg | aacacccccc | ccctggtgaa | gctgtggtac | 4920 |
| cagctggaga | aggagcccat | categgegee | gagaccttct | acgtggacgg | cgccgccaac | 4980 |
| cgcgagacca | agctgggcaa | ggccggctac | gtgaccgacc | gcggccgcca | gaaggtggtg | 5040 |

|            |            |            |            | ·          |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            | ccatccacct |            |      |
|            |            |            |            | acgccctggg |            |      |
|            |            |            |            | tcatcgagca |            |      |
|            |            |            |            | gcatcggcgg |            |      |
|            |            |            |            | tcctggacgg |            |      |
|            |            |            |            | ccatggccag |            |      |
|            |            |            |            | acaagtgcca |            |      |
|            |            |            |            | ggcagctgga |            |      |
|            |            |            |            | gcggctacat |            |      |
|            |            |            |            | tgctgaagct |            |      |
|            |            |            |            | tcaccagcac |            |      |
|            |            |            |            | tcccctacaa |            |      |
|            |            |            |            | tcatcggcca |            |      |
|            |            |            |            | tcatccacaa |            |      |
|            |            |            |            | tggacatcat |            |      |
| atccagacca | aggagctgca | gaagcagatc | accaagatcc | agaacttccg | cgtgtactac | 6000 |
|            |            |            |            | tgctgtggaa |            |      |
| gccgtggtga | tccaggacaa | cagcgacatc | aaggtggtgc | cccgccgcaa | ggccaagatc | 6120 |
| atccgcgact | acggcaagca | gatggccggc | gacgactgcg | tggccagccg | ccaggacgag | 6180 |
| gactaggaat | tcacgcgtcg | agcatgcatc | tagggcggcc | aattccgccc | ctctccctcc | 6240 |
| cccccccta  | acgttactgg | ccgaagccgc | ttggaataag | gccggtgtgc | gtttgtctat | 6300 |
|            |            |            |            | gggcccggaa |            |      |
|            |            |            |            | ccaaaggaat |            |      |
|            |            |            |            | gaagacaaac |            |      |
|            |            |            |            | ggtgcctctg |            |      |
|            |            |            |            | agtgccacgt |            |      |
|            |            |            |            | tcaacaaggg |            |      |
|            |            |            |            | ctcggtgcac |            |      |
|            |            |            |            | accacgggga |            |      |
|            |            |            |            | cctctagaac |            |      |
|            |            |            |            | ctggggcacc |            |      |
|            |            |            |            | gaccgtgtac |            |      |
|            |            |            |            | cgacgccaag |            |      |
|            |            |            |            | caccgacccc |            |      |
|            |            |            |            | gaagaacgac |            |      |
|            |            |            |            | gaagccctgc |            |      |
|            |            |            |            | cgacaccaac |            |      |
|            |            |            |            | ctgcagcttc |            |      |
|            |            |            |            | ctacaagctg |            |      |
|            |            |            |            | caacaccagc |            |      |
| aggestaces | caanningen | ttcgaggge  | teceestees | ctactgcgcc | cccaccacct | 7500 |
|            |            |            |            | gggcccctgc |            |      |
|            |            |            |            | cacccagctg |            |      |
|            |            |            |            | cttcgccgac |            |      |
|            |            |            |            | caccegeeee |            |      |
|            |            |            |            |            |            |      |
|            |            |            |            | caccaccggc |            |      |
|            |            |            |            | gtggaacgac |            |      |
|            |            |            |            | catcgtgttc |            |      |
|            |            |            |            | cggcggcgag |            |      |
|            |            |            |            | cacctggagc |            |      |
|            |            |            |            | ccgcatcaag |            |      |
|            |            |            |            | catcagcggc |            |      |
|            |            |            |            | cggcaacagc |            |      |
| gcgagatett | cogcetggge | ggcggcgaca | rgcgcgacaa | ctggcgcagc | gagetgtaca | 8280 |
| agcacaaggt | ggtgaagatc | gagcccctgg | gcgtggcccc | caccaaggcc | aagcgccgcg | 8340 |

| tggtgcagcg | cgagaagcgc | gccgtgggca | teggegeeet | gttcctgggc | ttcctgggcg | 8400  |
|------------|------------|------------|------------|------------|------------|-------|
| ccgccggcag | caccatgggc | gccgccagca | tgaccctgac | cgtgcaggcc | cgccagctgc | 8460  |
| tgagcggcat | cgtgcagcag | cagaacaacc | tgctgcgcgc | catcgaggcc | cagcagcacc | 8520  |
|            |            |            |            |            | gccgtggagc |       |
| gctacctgaa | ggaccagcag | ctgctgggca | tctggggctg | cagcggcaag | cttatctgca | 8640  |
| ccaccgccgt | gccctggaac | gccagctgga | gcaacaagag | cctggagcag | atctggaacc | 8700  |
| acaccacctg | gatggagtgg | gaccgcgaga | tcaacaacta | caccagcctg | atccacagcc | 8760  |
| tgatcgagga | gagccagaac | cagcaggaga | agaacgagca | ggagctgctg | gagctggaca | 8820  |
| agtgggccag | cctgtggaac | tggttcaaca | tcaccaactg | gctgtggtac | atcaagctgt | 8880  |
| tcatcatgat | cgtgggcggc | ctggtgggcc | tgcgcatcgt | gttcgccgtg | ctgagcatcg | 8940  |
| tgaaccgcgt | gcgccagggc | tacagccccc | tgagcttcca | gacccacctg | cccatccccc | 9000  |
| gcggccccga | ccgccccgag | ggcatcgagg | aggagggcgg | cgagcgcgac | cgcgaccgca | 9060  |
| gcatccgcct | ggtgaacggc | agcctggccc | tgatctggga | cgacctgcgc | agcctgtgcc | 9120  |
| tgttcagcta | ccaccgcctg | cgcgacctgc | tgctgatcgt | gacccgcatc | gtggagctgc | 9180  |
| tgggccgccg | cggctgggag | gccctgaagt | actggtggaa | cctgctgcag | tactggagcc | 9240  |
| aggagctgaa | gaacagcgcc | gtgagcctgc | tgaacgccac | cgccatcgcc | gtggccgagg | 9300  |
| gcaccgaccg | cgtgatcgag | gtggtgcagg | gcgcctgccg | cgccatccgc | cacatcccc  | 9360  |
| gccgcatccg | ccagggcctg | gagcgcatcc | tgctgtgagg | atccagatct | gctgtgcctt | 9420  |
| ctagttgcca | gccatctgtt | gtttgcccct | ccccgtgcc  | ttccttgacc | ctggaaggtg | 9480  |
| ccactcccac | tgtcctttcc | taataaaatg | aggaaattgc | atcgcattgt | ctgagtaggt | 9540  |
| gtcattctat | tctggggggt | ggggtggggc | aggacagcaa | gggggaggat | tgggaagaca | 9600  |
| atagcaggca | tgctggggat | gcggtgggct | ctatgggtac | ccaggtgctg | aagaattgac | 9660  |
| ccggttcctc | ctgggccaga | aagaagcagg | cacatcccct | tctctgtgac | acaccctgtc | 9720  |
| cacgcccctg | gttcttagtt | ccagccccac | tcataggaca | ctcatagctc | aggagggctc | 9780  |
| cgccttcaat | cccacccgct | aaagtacttg | gagcggtctc | tecetecete | atcagcccac | 9840  |
| caaaccaaac | ctagcctcca | agagtgggaa | gaaattaaag | caagataggc | tattaagtgc | 9900  |
| agagggagag | aaaatgcctc | caacatgtga | ggaagtaatg | agagaaatca | tagaatttta | 9960  |
| aggccatcat | ggccttaatc | ttccgcttcc | tegeteactg | actcgctgcg | ctcggtcgtt | 10020 |
|            |            |            |            |            | cacagaatca |       |
| ggggataacg | caggaaagaa | catgtgagca | aaaggccagc | aaaaggccag | gaaccgtaaa | 10140 |
| aaggccgcgt | tgctggcgtt | tttccatagg | ctccgcccc  | ctgacgagca | tcacaaaaat | 10200 |
| cgacgctcaa | gtcagaggtg | gcgaaacccg | acaggactat | aaagatacca | ggcgtttccc | 10260 |
| cctggaagct | ccctcgtgcg | ctctcctgtt | ccgaccctgc | cgcttaccgg | atacctgtcc | 10320 |
| gcctttctcc | cttcgggaag | cgtggcgctt | tctcatagct | cacgctgtag | gtatctcagt | 10380 |
| tcggtgtagg | tcgttcgctc | caagctgggc | tgtgtgcacg | aaccccccgt | tcagcccgac | 10440 |
| cgctgcgcct | tatccggtaa | ctatcgtctt | gagtccaacc | cggtaagaca | cgacttatcg | 10500 |
| ccactggcag | cagccactgg | taacaggatt | agcagagcga | ggtatgtagg | cggtgctaca | 10560 |
| gagttcttga | agtggtggcc | taactacggc | tacactagaa | gaacagtatt | tggtatctgc | 10620 |
|            |            |            |            |            | cggcaaacaa |       |
|            |            |            |            |            | cagaaaaaaa |       |
|            |            |            |            |            | gaacgaaaac |       |
|            |            |            |            |            | gatcctttta |       |
|            |            |            |            |            | gtctgacagt |       |
|            |            |            |            |            | ttcatccata |       |
|            |            |            |            |            | tgttgctgac |       |
|            |            |            |            |            | acggttgatg |       |
|            |            |            |            |            | cacggaacgg |       |
|            |            |            |            |            | tcgatttatt |       |
|            |            |            |            |            | aaccaattaa |       |
|            |            |            |            |            | ttcatatcag |       |
| _          |            |            |            |            | aactcaccga |       |
|            |            |            |            |            | cgtccaacat |       |
|            |            |            |            |            | aaatcaccat |       |
|            |            |            |            |            | cagacttgtt |       |
| caacaggcca | gccattacgc | tcgtcatcaa | aatcactcgc | atcaaccaaa | ccgttattca | 11640 |

```
ttcgtgattg cgcctgagcg agacgaaata cgcgatcgct gttaaaagga caattacaaa 11700 caggaatcga atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg 11760 aatcaggata ttcttctaat acctggaatg ctgtttccc gggggatcgca gtggtgagta 11820 accatgcatc atcaggagta cggataaaat gcttgatggt cggaagaggc ataaaatcccg 11880 tcagccagtt tagtctgacc atctcatctg taacatcatt ggcaacgcta cctttgccat 11940 gtttcagaaa caactctggc gcatcgggct tcccatacaa tcgatagatt gtcgcacctg 12000 attgcccgac attatcgcga gcccatttat acccatataa atcagcatc atgttggaat 12060 tcaatcgttat gtaagcagac agttttccc gttgaatatg gctcataaca ccccttgtat 12120 tactgttat gtaagcagac agttttattg ttcatgatga tatatttta tcttgtgcaa 12240 ttatcaggg ttattgcc atgagggat acatattga aagttttatg agacacaccg gccctttcgc tgacgtctaa gaaaccatta 12300 aaataggggt tccgcgaca tttccccgaa aagtgccacc tgacgtctaa gaaaccatta 12360 ttatcatgac attaacctat aaaaataggc gtatcacgag gccctttcgt c
```

<210> 166
<211> 9170
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 166

tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teageggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacqqtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegitta gigaacegie agategeeig gagaegeeat ceaegeigit tigaeeteea 960 tagaagacac egggacegat ecageeteeg eggeegggaa eggtgeattg gaacgeggat 1020 tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080 tettatgeat getataetgt ttttggettg gggeetatae acceeeqett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactecce 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae aeggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctqqtc ccatqcctcc 1500 agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcgqtagg qtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800

|            | gacataatag |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | gtcgtcgaca |            |            |            |            |      |
|            | aagatccgcc |            |            |            |            |      |
| cgtgtgggcc | agccgcgagc | tggagcgctt | cgccgtgaac | cccggcctgc | tggagaccag | 2040 |
| cgagggctgc | cgccagatcc | tgggccagct | gcagcccagc | ctgcagaccg | gcagcgagga | 2100 |
| gctgcgcagc | ctgtacaaca | ccgtggccac | cctgtactgc | gtgcaccagc | gcatcgagat | 2160 |
| caaggacacc | aaggaggccc | tggacaagat | cgaggaggag | cagaacaaga | gcaagaagaa | 2220 |
| ggcccagcag | gccgccgccg | acaccggcca | cagcaaccag | gtgagccaga | actaccccat | 2280 |
| cgtgcagaac | atccagggcc | agatggtgca | ccaggccatc | agcccccgca | ccctgaacgc | 2340 |
| ctgggtgaag | gtggtggagg | agaaggcctt | cagccccgag | gtgatcccca | tgttcagcgc | 2400 |
| cctgagcgag | ggcgccaccc | cccaggacct | gaacaccatg | ctgaacaccg | tgggcggcca | 2460 |
| ccaggccgcc | atgcagatgc | tgaaggagac | catcaacgag | gaggccgccg | agtgggaccg | 2520 |
|            | gtgcacgccg |            |            |            |            |      |
|            | ggcaccacca |            |            |            |            |      |
|            | gtgggcgaga |            |            |            |            |      |
|            | agccccacca |            |            |            |            |      |
|            | gaccgcttct |            |            |            |            |      |
|            | accgagaccc |            |            |            |            |      |
|            | ggccccgccg |            |            |            |            |      |
| cggccccggc | cacaaggccc | gcgtgctggc | cgaggccatg | agccaggtga | ccaacagcgc | 3000 |
|            | atgcagcgcg |            |            |            |            |      |
|            | gagggccaca |            |            |            |            |      |
|            | aaggagggcc |            |            |            |            |      |
|            | ttcccacaag |            |            |            |            |      |
|            | agaagagagc |            |            |            |            |      |
|            | agacaaggaa |            |            |            |            |      |
|            | acaataaaga |            |            |            |            |      |
|            | gtgctggagg |            |            |            |            |      |
|            | ggcttcatca |            |            |            |            |      |
|            | atcggcaccg |            |            |            |            |      |
|            | cagatcggct |            |            |            |            |      |
|            | aagcccggca |            |            |            |            |      |
|            | gccctggtgg |            |            |            |            |      |
|            | gagaacccct |            |            |            |            |      |
|            | aagctggtgg |            |            |            |            |      |
|            | ggcatccccc |            |            |            |            |      |
|            | gacgcctact |            |            |            |            |      |
|            | cccagcatca |            |            |            |            |      |
|            | tggaagggca |            |            |            |            |      |
|            | aagcagaacc |            |            |            |            |      |
|            | ctggagatcg |            |            |            |            |      |
|            | ggcttcacca |            |            |            |            |      |
|            | gagetgeace |            |            |            |            |      |
|            | accgtgaacg |            |            |            |            |      |
| gatctacgcc | ggcatcaagg | tacaccaact | gtgcaagctg | ctacacaaca | ccaaggccct | 4500 |
| gaccgaggtg | gtgcccctga | ccgaggaggc | cgagctggag | ctggccgaga | accocgagat | 4560 |
|            | cccgtgcacg |            |            |            |            |      |
|            | ggccagggcc |            |            |            |            |      |
|            | aagtacgccc |            |            |            |            |      |
|            | cagaagatcg |            |            |            |            |      |
| caagctgccg | atccagaagg | agacctggg  | gacctagtag | accgaotact | adcadaceac | 4860 |
|            | gagtgggagt |            |            |            |            |      |
|            | cccatcatcg |            |            |            |            |      |
|            | ggcaaggccg |            |            |            |            |      |
|            | accaaccaga |            |            |            |            |      |
| 340034666  |            |            |            |            |            |      |

| aaaaat aasa | atannanhaa | hanaaanaa  |            |            | h          | F1.60 |
|-------------|------------|------------|------------|------------|------------|-------|
|             |            |            |            |            | tccaggccca |       |
|             |            |            |            |            | tcaagaagga |       |
|             |            |            |            |            | agcaggtgga |       |
|             |            |            |            |            | acaaggccca |       |
| ggaggagcac  | gagaagtacc | acagcaactg | gcgcgccatg | gccagcgact | tcaacctgcc | 5400  |
| ccccgtggtg  | gccaaggaga | tcgtggccag | ctgcgacaag | tgccagctga | agggcgaggc | 5460  |
| catgcacggc  | caggtggact | gcagccccgg | catctggcag | ctggcatgca | cccacctgga | 5520  |
|             |            |            |            |            | ccgaggtgat |       |
|             |            |            |            |            | gccgctggcc |       |
|             |            |            |            |            | tgaaggccgc |       |
|             |            |            |            |            | agagccaggg |       |
|             |            |            |            |            | gcgaccaggc |       |
|             |            |            |            |            | agcgcaaggg |       |
|             |            |            |            |            | ccgacatcca |       |
|             |            |            |            |            |            |       |
|             |            |            |            |            | actaccgcga |       |
|             |            |            |            |            | agggcgccgt |       |
|             |            |            |            |            | agatcatccg |       |
|             |            |            |            |            | acgaggacta |       |
|             |            |            |            |            | ccccgtgcct |       |
|             |            |            |            |            | ggaaattgca |       |
|             |            |            |            |            | gcacagcaag |       |
|             |            |            |            |            | tatgggtacc |       |
| caggtgctga  | agaattgacc | cggttcctcc | tgggccagaa | agaagcaggc | acatcccctt | 6480  |
| ctctgtgaca  | caccctgtcc | acgcccctgg | ttcttagttc | cagccccact | cataggacac | 6540  |
| tcatagctca  | ggagggctcc | gccttcaatc | ccacccgcta | aagtacttgg | agcggtctct | 6600  |
|             |            |            |            |            | aaattaaagc |       |
|             |            |            |            |            | gaagtaatga |       |
|             |            |            |            |            | cggtcgttcg |       |
|             |            |            |            |            | cagaatcagg |       |
|             |            |            |            |            | accgtaaaaa |       |
|             |            |            |            |            | acaaaaatcg |       |
|             |            |            |            |            | cgtttcccc  |       |
|             |            |            |            |            | acctgtccgc |       |
|             |            |            |            |            | atctcagttc |       |
|             |            |            |            |            |            |       |
|             |            |            |            |            | agecegaceg |       |
|             |            |            |            |            | acttatcgcc |       |
|             |            |            |            |            | gtgctacaga |       |
|             |            |            |            |            | gtatctgcgc |       |
|             |            |            |            |            | gcaaacaaac |       |
|             |            |            |            |            | gaaaaaaagg |       |
|             |            |            |            |            | acgaaaactc |       |
|             |            |            |            |            | tccttttaaa |       |
|             |            |            |            |            | ctgacagtta |       |
|             |            |            |            |            | catccatagt |       |
|             |            |            |            |            | gttgctgact |       |
|             |            |            |            |            | cggttgatga |       |
| gagctttgtt  | gtaggtggac | cagttggtga | ttttgaactt | ttgctttgcc | acggaacggt | 7920  |
| ctgcgttgtc  | gggaagatgc | gtgatctgat | ccttcaactc | agcaaaagtt | cgatttattc | 7980  |
| aacaaagccg  | ccgtcccgtc | aagtcagcgt | aatgctctgc | cagtgttaca | accaattaac | 8040  |
|             |            |            |            |            | tcatatcagg |       |
|             |            |            |            |            | actcaccgag |       |
|             |            |            |            |            | gtccaacatc |       |
|             |            |            |            |            | aatcaccatg |       |
|             |            |            |            |            | agacttgttc |       |
| aacaggggag  | ccattacact | cotcatcaaa | atcactcoca | tcaaccaaac | cgttattcat | 8400  |
|             |            | •          |            | =          |            |       |

```
tegtgattge geetgagega gaegaaatae gegategetg ttaaaaggae aattacaaae 8460 aggaategaa tgeaacegge geaggaacae tgeeagegea teaacaatat ttteaectga 8520 ateaggatat tettetaata eetggaatge tgttteeeg gggategeag tggtgagtaa 8580 ceatgeatea teaaggagea teeteategg ggataaaatg etgttgggee ggaagaggea taaatteegt 8640 cagecagttt agtetgacea teteatetgt aacateattg geaaegetae etttgeeatg 8700 ttteagaaae aactetggeg eateggget eecaattaaa teageageate teteatega 8760 ttgeeegaca ttategegag eecaattata eecaataaa teageatee tgttggaatt 8820 taateggge etegageaag aegttteeeg ttgaatatgg eteataaae eecettgatt 8880 aetgttatg taageagaa getttatgt teatgatgat atattttat ettgtgeaat 8940 gtaacateag agattttgag acacaaegtg gettteeee eeceecatt attgaageat 9000 ttateagggt tattgeea tteeegaaa agtgeeaeet gaegtetaag aaaacaataa 9120 tateatgaea ttaacetata aaaataageg tateaegagg eeettteegte 9170
```

<210> 167
<211> 9280
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing SIV genes

<400> 167

tegegegett eggegatgae ggegaaaace teegacacat geageteeeg gagaeggeea 60 cagettqtet qtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg qtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgaccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatac acceeegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagaetgae aeggaetetg tatttttaca 1320 qqatqggqtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcqaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte coatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggeeag aettaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 qcaqaagaag atgcaggcag ctgagttgtt gtattctqat aagagtcaga ggtaactccc 1740 gttgeggtge tgttaacggt ggagggeagt gtagtetgag cagtactcgt tgctgccgcg 1800

|            | ~~~~~~~    | atasasas   |            | teatteast. | aaatatttta | 1060 |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | gggtcttttc |      |
| tgeagteace | gregregacg | cegecaceat | gggcgcgcgc | aacagegege | tgagcggcaa | 1000 |
| gaaggccgac | gagetggaga | agateegeet | gegeeegaae | ggcaagaaga | agtacatgct | 1980 |
| gaagcacgtg | araraaacca | ccaacgagct | ggaccgcttc | agecraacea | agagcctgct | 2040 |
| ggagaacaag | gagggctgcc | agaagateet | gagegegetg | gcgccgccgg | tgcccaccgg | 2100 |
| cagcgagaac | ctgaagagcc | tgtacaacac | cgtgtgcgtg | atctggtgca | tccacgccga | 2160 |
| ggagaaggtg | aagcacaccg | aggaggccaa | gcagatcgtg | cagcgccacc | tggtggtgga | 2220 |
| gaccggcacc | accgagacca | tgcccaagac | cagccgcccg | accgcgccca | gcagcggacg | 2280 |
| cggcggcaac | tacccggtgc | agcagatcgg | cggcaactac | gtgcacctgc | cgctgagccc | 2340 |
|            |            |            |            |            | ccgaggtggt |      |
| gcccggcttc | caggccctga | gcgagggctg | cacgccctac | gacatcaacc | agatgctgaa | 2460 |
| ctgcgtgggc | gaccaccagg | ccgccatgca | gatcatccgc | gacatcatca | acgaggaggc | 2520 |
| cgccgactgg | gacctgcagc | acccgcagcc | cgcgccgcag | cagggccagc | tgcgcgagcc | 2580 |
|            |            |            |            |            | agtggatgta |      |
|            |            |            |            |            | agctgggcct |      |
|            |            |            |            |            | agggcccgaa |      |
|            |            |            |            |            | agcagaccga |      |
|            |            |            |            |            | acccggactg |      |
|            |            |            |            |            | tgaccgcctg |      |
|            |            |            |            |            | tgaaggaggc |      |
|            |            |            |            |            | gcaagcccat |      |
|            |            |            |            |            | cgccgcgccg |      |
|            |            |            |            |            | cggaccgcca |      |
|            |            |            |            |            | gctcaagtgc |      |
|            |            |            |            |            | ctaaagaact |      |
|            |            |            |            |            | aagccttaca |      |
|            |            |            |            |            | cagtagtcac |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | acgacagcat |      |
|            |            |            |            |            | gcatcggcgg |      |
|            |            |            |            |            | agegeateaa |      |
|            |            |            |            |            | tgctgaccgc |      |
|            |            |            |            |            | tggccctgaa |      |
|            |            |            |            |            | agatcgtggc |      |
|            |            |            |            |            | cccacccac  |      |
|            |            |            |            |            | agtggcgcat |      |
|            |            |            |            |            | tgcagctggg |      |
|            |            |            |            |            | acatcggcga |      |
|            |            |            |            |            | tcaccctgcc |      |
|            |            |            |            |            | cccagggctg |      |
|            |            |            |            |            | ccttccgcaa |      |
|            |            |            |            |            | ccagcgaccg |      |
|            |            |            |            |            | acagcatcgg |      |
|            |            |            |            |            | tgggctacga |      |
|            |            |            |            |            | agacctggac |      |
|            |            |            |            |            | tctaccccgg |      |
|            |            |            |            |            | ccgaggaggt |      |
| gcagtggacc | gagatggccg | aggccgagta | cgaggagaac | aagatcatcc | tgagccagga | 4620 |
|            |            |            |            |            | tcaagagcca |      |
|            |            |            |            |            | tgggcaagtt |      |
|            |            |            |            |            | tgatccagaa |      |
| gatcggcaag | gaggccatcg | tgatctgggg | ccaggtgccc | aagttccacc | tgcccgtgga | 4860 |
| gaaggacgtg | tgggagcagt | ggtggaccga | ctactggcag | gtgacctgga | tacccgagtg | 4920 |
|            |            |            |            |            | aggaccccat |      |
|            |            |            |            |            | aggagggcaa |      |
|            |            |            |            |            | agaccaccaa |      |
|            | -          |            |            |            |            |      |

| ccagcaggcc | gagetegagg   | ccttcctqat   | qqccctgacc   | gacagcggcc    | ccaaggccaa   | 5160         |
|------------|--------------|--------------|--------------|---------------|--------------|--------------|
| catcatcata | gacagecagt   | acqtqatqqq   | catcatcacc   | ggctgcccca    | ccgagagcga   | 5220         |
| gagccgcctg | gtgaaccaga   | tcatcgagga   | gatgatcaag   | aagagcgaga    | tctacgtggc   | 5280         |
| ctagataccc | acccacaagg   | gcatcggcgg   | caaccaggag   | atcgaccacc    | tggtgagcca   | 5340         |
| aggcatccgc | caggtgctgt   | tectggagaa   | gategageee   | gcccaggagg    | agcacgacaa   | 5400         |
| gtaccacage | aacgtgaagg   | agctggtgtt   | caaqttcggc   | ctgcccgca     | tcgtggcccg   | 5460         |
| ccagatcgtg | gacacctgcg   | acaaqtqcca   | ccaqaaqqqc   | gaggccatcc    | acggccaggc   | 5520         |
| caacagcgac | chagacacct   | ggcagatggc   | ctqcacccac   | ctggagggca    | agatcatcat   | 558 <b>0</b> |
| cataaccata | cacatageca   | gcggcttcat   | cqaqqccqaq   | gtgatcccgc    | aggagaccgg   | 5640         |
| caccadacc  | accetattee   | tactaaaact   | aaccaaccac   | tggcccatca    | cccacctgca   | 5700         |
| caccacaac  | gacaccaact   | tcaccaacca   | ggaggtgaag   | atggtggcct    | qqtqqqccqq   | 5760         |
| categacaac | accttcaca    | toccctacaa   | cccacadadc   | cagggcgtgg    | tagaagccat   | 5820         |
| gaagagagag | ctcaacaacc   | agatroacco   | catecocoao   | caggccaaca    | gcgtggagac   | 5880         |
| gaaccaccac | atagacatac   | actoratoiaa  | chtcaagcgc   | cgcggcggca    | toggogacat   | 5940         |
| categugetg | arggregege   | tcaacatgat   | caccaccaaa   | caggagatcc    | agttccagca   | 6000         |
| gaegeeegee | agegeetga    | agaacttcca   | catatactac   | cgcgagggcc    | gcgaccagct   | 6060         |
| gagcaagaac | agcaagccca   | tactatage    | caccasacac   | gccgtgatcc    | taaaaataaa   | 6120         |
| grggaagggc | eceggegage   | coccoccoc    | gggcgagggc   | atcaaggact    | earraged and | 6180         |
| caccgacacc | aaggeggege   | geegeegeaa   | ggccaagacc   | accaaggace    | acceptaceta  | 6240         |
| caaggaggtg | gacagcagca   | gccacatgga   | ggacaccggc   | gaggcccgcg    | tagttgccag   | 6300         |
| agaattcgtt | taaacacgeg   | b-set        | tectagaceeg  | ctgtgccttc    | cacteceact   | 6360         |
| ccatctgttg | tttgccctc    | eccegtgeet   | tegestate    | tggaaggtgc    | teattetatt   | 6420         |
| greettreet | aataaaatga   | ggaaactgca   | regeattgtt   | tgagtaggtg    | tancannet    | 6480         |
| ctggggggtg | gggtggggca   | ggacagcaag   | ggggaggacc   | gggaagacaa    | cagcaggcac   | 6540         |
| gctggggatg | cggtgggctc   | tatgggtacc   | cagguguuga   | agaattgacc    | eggeteetee   | 6600         |
| tgggccagaa | agaagcaggc   | acateceett   | ctctgtgaca   | caccctgtcc    | aegeeeetgg   | 6600         |
| ttcttagttc | cagccccact   | cataggacac   | tcatagetca   | ggagggctcc    | gccttcaatc   | 6550         |
| ccacccgcta | aagtacttgg   | ageggtetet   | ccctccctca   | tcagcccacc    | aaaccaaacc   | 6720         |
| tagcctccaa | gagtgggaag   | aaattaaagc   | aagataggct   | attaagtgca    | gagggagaga   | 6780         |
| aaatgcctcc | aacatgtgag   | gaagtaatga   | gagaaatcat   | agaattttaa    | ggccatcatg   | 6040         |
| gccttaatct | teegetteet   | cgctcactga   | ctegetgege   | teggtegtte    | ggetgeggeg   | 6900         |
| agcggtatca | gctcactcaa   | aggcggtaat   | acggttatcc   | acagaatcag    | gggataacgc   | 7020         |
| aggaaagaac | atgtgagcaa   | aaggccagca   | aaaggccagg   | aaccgtaaaa    | aggeegegee   | 7020         |
| gctggcgttt | ttccataggc   | reegeceee    | tgacgagcat   | cacaaaaatc    | gacgcccaag   | 7080         |
| tcagaggtgg | cgaaacccga   | caggactata   | aagataccag   | gcgtttcccc    | etggaagete   | 7200         |
| cctcgtgcgc | tetectgtte   | cgaccctgcc   | gcttaccgga   | tacctgtccg    | CCLLCCCC     | 7200         |
| ttcgggaagc | gtggcgcttt   | ctcatagete   | acgetgtagg   | tatctcagtt    | cggtgtaggt   | 7260         |
| cgttcgctcc | aagctgggct   | gtgtgcacga   | acccccgtt    | cagecegace    | gctgcgcctt   | 7320         |
| atccggtaac | tatcgtcttg   | agtccaaccc   | ggtaagacac   | gacttatcgc    | cactggcagc   | 7380         |
| agccactggt | aacaggatta   | gcagagcgag   | gtatgtaggo   | ggtgctacag    | agttcttgaa   | 7440         |
| gtggtggcct | aactacggct   | acactagaag   | aacagtattt   | ggtatctgcg    | ctctgctgaa   | 7500         |
| gccagttacc | ttcggaaaaa   | gagttggtag   | ctcttgatcc   | ggcaaacaaa    | ccaccgctgg   | 7560         |
| tagcggtggt | ttttttgttt   | gcaagcagca   | gattacgcgc   | : agaaaaaaag  | gateteaaga   | 7620         |
| agatcctttg | atcttttcta   | cggggtctga   | cgctcagtgg   | , aacgaaaact  | cacgttaagg   | 7680         |
| gattttggto | : atgagattat | caaaaaggat   | . cttcacctag | , atccttttaa  | attaaaaatg   | 7740         |
| aagttttaaa | tcaatctaaa   | gtatatatga   | gtaaacttgg   | tctgacagtt    | accaatgctt   | 7800         |
| aatcagtgag | gcacctatct   | cagcgatctg   | , tctatttcgt | tcatccatag    | ttgcctgact   | 7860         |
| cggggggggg | , gggcgctgag | gtctgcctcg   | , tgaagaaggt | gttgctgact    | cataccaggc   | 7920         |
| ctgaatcgcc | : ccatcatcca | gccagaaagt   | : gagggagcca | a cggttgatga  | gagctttgtt   | 7980         |
| qtaggtggad | : cagttggtga | ttttgaactt   | : ttgctttgcc | : acggaacggt  | ctgcgttgtc   | 8040         |
| qqgaagatgo | gtgatctgat   | : ccttcaactc | : agcaaaagtt | : cgatttattc  | aacaaagccg   | 8100         |
| ccqtcccqt  | aaqtcaqcqt   | : aatgctctgc | : cagtgttaca | a accaattaac  | caattctgat   | 8160         |
| tagaaaaact | catcgagcat   | caaatgaaac   | : tgcaatttat | : 'tcatatcagg | attatcaata   | 8220         |
| ccatatttt  | gaaaaagccg   | tttctgtaat   | : gaaggagaaa | a actcaccgag  | gcagttccat   | 8280         |
| aggatggcaa | a gatectggta | toggtotgog   | g attecgacto | gtccaacatc    | aatacaacct   | 8340         |
| attaatttc  | cctcgtcaaa   | aataaggtta   | a tcaagtgaga | a aatcaccatg  | agtgacgact   | 8400         |
|            | _            |              |              |               |              |              |

```
quatcoqqtq aqaatqqcaa aagcttatgc atttctttcc agacttgttc aacaggccag 8460
ccattacgct cgtcatcaaa atcactcgca tcaaccaaac cgttattcat tcgtgattgc 8520
gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac aattacaaac aggaatcgaa 8580
tgcaaccggc gcaggaacac tgccagcgca tcaacaatat tttcacctga atcaggatat 8640
tettetaata cetggaatge tgtttteeeg gggategeag tggtgagtaa ceatgeatea 8700
traggagtar ggataaaatg cttgatggtr ggaagaggra taaattregt ragcragttt 8760
agtotgacca totcatotgt aacatoattg goaacgotac otttgccatg totcagaaac 8820
aactotggcg categggett cocatacaat cgatagattg tegcacetga ttgccegaca 8880
ttatcgcgag cccatttata cccatataaa tcagcatcca tgttggaatt taatcgcggc 8940
ctcgagcaag acgtttcccg ttgaatatgg ctcataacac cccttgtatt actgtttatg 9000
taagcagaca gttttattgt tcatgatgat atatttttat cttgtgcaat gtaacatcag 9060
agattttgag acacaacgtg gctttccccc ccccccatt attgaagcat ttatcagggt 9120
tattqtctca tqaqcqqata catatttqaa tgtatttaga aaaataaaca aataggggtt 9180
ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca 9240
                                                                  9280
ttaacctata aaaataggcg tatcacgagg ccctttcgtc
```

<210> 168
<211> 9166
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 168 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggeeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 qqqqtcatta qttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cqtcaatqqq agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcqttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960 taqaagacac egggacegat ecageeteeg eggeegggaa eggtgeattg gaacgeggat 1020 tecceptace aagagtgacg taagtacege ctatagacte tataggeaca eccetttgge 1080 tettatgcat getatactgt ttttggettg gggeetatac acceecgett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct teteoggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 aggggeteat ggtegetegg cageteettg etectaacag tggaggeeag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

|            |            |            |            |            | ggtaactccc |      |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            | tgctgccgcg |      |
|            |            |            |            |            | gggtcttttc |      |
|            |            |            |            |            | gcaagctgga |      |
| cgcctgggag | cgcatcaagc | tgaagcccgg | cggcaagaag | cactacatga | tgaagcacct | 1980 |
| ggtgtgggcc | agccgcgagc | tggagcgctt | cgccctggac | cccggcctgc | tggagaccag | 2040 |
| cgagggctgc | aagcagatca | tgaagcagct | gcagcccgcc | ctgcagaccg | gcaccaagga | 2100 |
|            |            |            |            |            | agatcgacgt |      |
|            |            |            |            |            | gccagcagaa |      |
|            |            |            |            |            | ccatcgtgca |      |
|            |            |            |            |            | acgcctgggt |      |
|            |            |            |            |            | ccgccctgag |      |
|            |            |            |            |            | gccaccaggc |      |
|            |            |            |            |            | accgcctgca |      |
|            |            |            |            |            | gcagcgacat |      |
|            |            |            |            |            | accccccgt  |      |
|            |            |            |            |            | tcgtgcgcat |      |
|            |            |            |            |            | tccgcgacta |      |
|            |            |            |            |            | tgaagaactg |      |
|            |            |            |            |            | teetgegege |      |
|            |            |            |            |            | tgggcggccc |      |
|            |            |            |            |            | ccaacatcat |      |
|            |            |            |            |            | actgcggcaa |      |
|            |            |            |            |            |            |      |
| ggagggccac | ecggeeegea | accyccycyc | ggaggagag  | aagggccgcc | ggaagtgcgg | 3120 |
|            |            |            |            |            | tagggaagat |      |
|            |            |            |            |            | agccaacagc |      |
|            |            |            |            |            | agaagcagga |      |
|            |            |            |            |            | gcagcgaccc |      |
|            |            |            |            |            | cggcgccgac |      |
|            |            |            |            |            | gatcggcggc |      |
|            |            |            |            |            | ctgcggccac |      |
|            |            |            |            |            | ccgcaacctg |      |
|            |            |            |            |            | cgtgcccgtg |      |
|            |            |            |            |            | cgaggagaag |      |
|            |            |            |            |            | cagcaagatc |      |
|            |            |            |            |            | cagcaccaag |      |
|            |            |            |            |            | ctgggaggtg |      |
|            |            |            |            |            | cgtgctggac |      |
|            |            |            |            |            | caccgccttc |      |
|            |            |            |            |            | cgtgctgccc |      |
|            |            |            |            |            | cctggagccc |      |
|            |            |            |            |            | gtacgtgggc |      |
|            |            |            |            |            | gcacctgctg |      |
| cgctggggct | tcaccacccc | cgacaagaag | caccagaagg | agcccccctt | cctgtggatg | 4320 |
|            |            |            |            |            | cgagaaggac |      |
|            |            |            |            |            | cagccagatc |      |
|            |            |            |            |            | ggccctgacc |      |
|            |            |            |            |            | cgagatcctg |      |
|            |            |            |            |            | cgagatccag |      |
|            |            |            |            |            | gaacctgaag |      |
|            |            |            |            |            | gctgaccgag |      |
|            |            |            |            |            | caagttcaag |      |
|            |            |            |            |            | ggccacctgg |      |
|            |            |            |            |            | ccagctggag |      |
| aaggagccca | tcatcggcgc | cgagaccttc | tacgtggacg | gcgccgccaa | ccgcgagacc | 4980 |

aagctgggca aggccggcta cgtgaccgac cgcggccgcc agaaggtggt gcccctgacc 5040 gacaccacca accagaagac cgagctgcag gccatccacc tggccctgca ggacagcggc 5100 ctggaggtga acatcgtgac cgacagccag tacgccctgg gcatcatcca ggcccagccc 5160 gacaagagcg agagcgagct ggtgagccag atcatcgagc agctgatcaa gaaggagaag 5220 gtgtacetgg cetgggtgce egeceacaag ggcateggeg geaacgagea ggtggaegge 5280 ctggtgagcg ccggcatccg caaggtgctg ttcctggacg gcatcgacaa ggcccaggag 5340 gagcacgaga agtaccacag caactggege gccatggeca gegacttcaa cetgeecece 5400 gtggtggcca aggagatcgt ggccagctgc gacaagtgcc agctgaaggg cgaggccatg 5460 cacggccagg tggactgcag ccccggcatc tggcagctgg catgcaccca cctggagggc 5520 aaggtgatee tggtggeegt geaegtggee ageggetaea tegaggeega ggtgateeee 5580 gccgagaccg gccaggagac cgcctacttc ctgctgaagc tggccggccg ctggcccgtg 5640 aagaccgtgc acaccgacaa cggcagcaac ttcaccagca ccaccgtgaa ggccgcctgc 5700 tggtgggccg gcatcaagca ggagttcggc atcccctaca acccccagag ccagggcgtg 5760 atcgagagca tgaacaagga gctgaagaag atcatcggcc aggtgcgcga ccaggccgag 5820 cacctgaaga ccgccgtgca gatggccgtg ttcatccaca acttcaagcg caagggcggc 5880 ateggegget acagegeegg egagegeate gtggacatea tegecacega cateeagace 5940 aaggagetge agaageagat caccaagate cagaaettee gegtgtaeta eegegacage 6000 egegaeeeeg tgtggaaggg ceeegeeaag etgetgtgga agggegaggg egeegtggtg 6060 atccaqqaca acaqcqacat caagqtqqtq ccccqccqca aggccaagat catccqcqac 6120 tacqqcaaqc aqatqqccqq cqacqactqc qtqqccaqcc gccaqqacqa qqactaqqaa 6180 ttottgctgt gccttctagt tgccagccat ctgttgtttg cccctcccc gtgccttcct 6240 tgaccetgga aggtgecact eccaetgtee ttteetaata aaatgaggaa attgeatege 6300 aggattggga agacaatagc aggcatgctg gggatgcggt gggctctatg ggtacccagg 6420 tgctgaagaa ttgacccggt tcctcctggg ccagaaagaa gcaggcacat ccccttctct 6480 gtgacacacc ctgtccacgc ccctggttct tagttccagc cccactcata ggacactcat 6540 ageteaggag ggeteegeet teaateeeac eegetaaagt aettggageg gteteteeet 6600 ccctcatcag cccaccaaac caaacctagc ctccaagagt gggaagaaat taaagcaaga 6660 taggctatta agtgcagagg gagagaaaat gcctccaaca tgtgaggaag taatgagaga 6720 aatcatagaa tttcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg 6780 cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat 6840 aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 6900 gegttgetgg egttttteca taggeteege ecceetgaeg agcateacaa aaategaege 6960 tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 7020 ageteceteg tgegetetec tgtteegace etgeegetta eeggatacet gteegeettt 7080 ctecettegg gaagegtgge gettteteaa tgeteaeget gtaggtatet eagtteggtg 7140 taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 7200 geettateeg gtaactateg tettgagtee aacceggtaa gacacgactt ategecactg 7260 qcaqcaqcca ctqqtaacaq qattaqcaqa qcgaggtatg taggcggtgc tacagagttc 7320 ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg 7380 ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 7440 gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct 7500 caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt 7560 taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa 7620 aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa 7680 tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc 7740 tgactccggg ggggggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata 7800 ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc 7860 tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc 7920 gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca 7980 aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 8040 tetgattaga aaaaeteate gageateaaa tgaaaetgea atttatteat ateaggatta 8100 tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 8160 ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 8220 caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 8280

```
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgtcaaca 8340 ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 8400 gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt accaaacagga 8460 atcgaatgca accggcgcag gaacactgcc agcgcatcaa caatattttc acctgaatca 8520 ggatattctt ctaatacctg gaatgctgtt ttcccgggga tcgcagtggt gagtaaccat 8580 gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 8640 cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgttc 8700 agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 8760 ccgacattat cgcgagccca tttataccca tataaaacag catcatgtt ggaatttaat 8820 cgcggcctcg agcaagacgt ttcccgttga atatggctca taacacccct tgtattactg 8880 ttatgtaag cagacagtt tattgttcat gatgatata ttttatcttg tgcaatgtaa 8940 cagggttatt gtctcatgag cggatacata ttttgaatgta tttaagaaaa taaacaacta 9060 ggggttccgc gcacatttc ccgaaaaagt ccacctgacg tccacctgacg ttcaagaaac cattattac 9120 atgacattaa cctataaaaa cctataaaaa taagcgccct ttcgtc 9166
```

<210> 169
<211> 9169
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

## <400> 169

tegegegtht eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 eccgcetggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtetcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegthta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1080 tettatgeat getatactgt tittggettg gggeetatae acccegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae acggaetetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 egeagttttt attaaacata gegtgggate tecaegegaa tetegggtae gtgtteegga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ecatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag acttaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

| gcagaagaag |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| gttgcggtgc |            |            |            |            |            |      |
| cgcgccacca |            |            |            |            |            |      |
| tgcagtcacc |            |            |            |            |            |      |
| cgcctgggag |            |            |            |            |            |      |
| gatctgggcc |            |            |            |            |            |      |
| cgacggctgc | cagcagatca | tggagaagct | gcagcccgcc | ctgaagaccg | gcaccgagga | 2100 |
| gatcaagagc | ctgttcaaca | ccgtggccac | cctgtactgc | gtgcaccagc | gcatcgacgt | 2160 |
|            |            |            | cgaggagatc |            |            |      |
| gacccagcag | gccgccgccg | acaccggcaa | cagcagcaac | gtgagccaga | actaccccat | 2280 |
| cgtgcagaac | gcccagggcc | agatgatcca | ccagagcctg | agcccccgca | ccctgaacgc | 2340 |
| ctgggtgaag | gtgatcgagg | agaaggcctt | cagccccgag | gtgatcccca | tgttcagcgc | 2400 |
| cctgagcgag | ggcgccaccc | cccaggacct | gaacatgatg | ctgaacatcg | tgggcggcca | 2460 |
|            |            |            | catcaacgag |            |            |      |
| cctgcacccc | gtgcacgccg | gccccatccc | ccccggccag | atgcgcgagc | cccgcggcag | 2580 |
| cgacatcgcc | ggcaccacca | gcacccccca | ggagcagatc | ggctggatga | ccggcaaccc | 2640 |
|            |            |            | ctggatcatc |            |            |      |
| gcgcatgtac | agccccgtga | gcatcctgga | catcaagcag | ggccccaagg | agcccttccg | 2760 |
| cgactacgtg | gaccgcttct | tcaagacctt | ccgcgccgag | caggccaccc | aggacgtgaa | 2820 |
| gaactggatg | accgacaccc | tgctggtgca | gaacgccaac | cccgactgca | aggccatcct | 2880 |
| gegegeeetg | ggccccggcg | ccaccctgga | ggagatgatg | accgcctgcc | agggcgtggg | 2940 |
|            |            |            | cgaggccatg |            |            |      |
| catcatgatg | cagcgcggca | acttcaaggg | ccagaagcgc | atcaagtgct | tcaactgcgg | 3060 |
|            |            |            | cgccctgcgc |            |            |      |
|            |            |            | caccgagcga |            |            |      |
| gatctggcct | tcccacaagg | gaaggccagg | gaattttctt | cagagcagac | cagagccaac | 3240 |
|            |            |            | ggaagagaca |            |            |      |
| ggagccgata | gacaaggaac | tgtatccttt | agcttccctc | agatcactct | ttggcagcga | 3360 |
| cccctcgtca | caataaagat | agggggccag | ctgaaggagg | cccttctaga | caccggcgcc | 3420 |
| gacgacaccg | tgctggagga | gatgaacctg | cccggccgct | ggaagcccaa | gatgatcggc | 3480 |
|            |            |            | tacgaccaga |            |            |      |
|            |            |            | cccacccccg |            |            |      |
|            |            |            | ttccccatca |            |            |      |
|            |            |            | aaggtgaagc |            |            |      |
|            |            |            | gagatggaga |            |            |      |
|            |            |            | gtgttcgcca |            |            |      |
|            |            |            | ctgaacaagc |            |            |      |
|            |            |            | ctgaagcaga |            |            |      |
|            |            |            | ctggacaagg |            |            |      |
|            |            |            | cccggcatcc |            |            |      |
|            |            |            | ttccagtgca |            |            |      |
| cccttccgca | agcagaaccc | cgacatcgtg | atctaccagt | acatggacca | cctgtacgtg | 4200 |
| ggcagcgacc | tggagatcgg | ccagcaccgc | accaagatcg | aggagctgcg | ccagcacctg | 4260 |
|            |            |            | aagcaccaga |            |            |      |
| atgggctacg | agctgcaccc | cgacaagtgg | accgtgcagc | ccatcgtgct | gcccgagaag | 4380 |
| gacagetgga | ccgtgaacga | catccagaag | ctggtgggca | agctgaactg | ggccagccag | 4440 |
|            |            |            | tgcaagctgc |            |            |      |
|            |            |            | gagctggagc |            |            |      |
|            |            |            | gaccccagca |            |            |      |
|            |            |            | cagatctacc |            |            |      |
|            |            |            | gcccacacca |            |            |      |
|            |            |            | atcgtgatct |            |            |      |
|            |            |            | gcctggtgga |            |            |      |
|            |            |            | cccccctgg  |            |            |      |
|            |            |            | ttctacgtgg |            |            |      |
|            |            |            |            |            |            |      |

| acceanctor | acaaaaccaa | ctacqtqacc | gaccgcggcc | accadaadat  | ggtgcccctg | 5040 |
|------------|------------|------------|------------|-------------|------------|------|
| accaagecgg | ccaaccadaa | gaccaageta | caddccatcc | acctggccct  | gcaggacagc | 5100 |
| accededea  | toaacatcot | gaccgacage | cagtacgccc | toggcatcat  | ccaggcccag | 5160 |
| 2200032423 | acasasacas | actaataaac | cagatcatcg | agcagctgat  | caagaaggag | 5220 |
| aaggtgtacc | taacctaaat | acccacces  | aagggcatcg | gcggcaacga  | gcaggtggac | 5280 |
| aaggegeace | acaccaacat | ccacaaata  | chahheetaa | acggcatcga  | caaggcccag | 5340 |
| ggcctggtga | agaagtacca | cagcaactag | cacaccataa | ccagcgactt  | caacctgccc | 5400 |
| gaggagcacg | ccaaggagat | cataaccaac | tocoacaadt | accaactass  | gggcgaggcc | 5460 |
| atagagaga  | agatagagac | caccacac   | atctggcage | tagcatacac  | ccacctggag | 5520 |
| acgeacggee | tectagtage | catacacata | accaggage  | acaticgagge | cgaggtgatc | 5580 |
| ggcaaggcga | ccccggcggc | gegeacgeg  | ttcctcctca | acctagagge  | ccgctggccc | 5640 |
| cccgccgaga | tagacagga  | gaccgcccac | aacttcacca | acaccaccat  | gaaggccgcc | 5700 |
| grgaagaccg | cgcacaccga | ggaggagthe | ggcategggt | acaaccccca  | gagccagggc | 5760 |
| cgetggtggg | ceggeaceaa | gcaggageee | aggatcatca | accadeceed  | cgaccaggcc | 5820 |
| gtgattgaga | gcatgaacaa | ggagetgaag | atattatta  | acaacttcaa  | acacaaaaaa | 5880 |
| gagcacctga | agacegeege | geagatggee | ataataaaaa | testegges   | gcgcaagggc | 5940 |
| ggcatcggcg | gctacagcgc | eggegagege | ategeggaca | tagaggtata  | cgacatccag | 6000 |
| accaaggagc | tgcagaagca | gatcaccaag | acccagaact | ceegegegea  | ctaccgcgac | 6060 |
| agccgcgacc | ccgtgtggaa | gggcccgcc  | aagetgetgt | ggaagggcga  | gggcgccgtg | 6120 |
| gtgatccagg | acaacagcga | catcaaggtg | grgeeeegee | geaaggeeaa  | gatcatccgc | 0120 |
| gactacggca | agcagatggc | cggcgacgac | tgcgtggcca | geegeeagga  | cgaggactag | 6180 |
| gaattcttgc | tgtgccttct | agttgccagc | catctgttgt | ttgccctcc   | cccgtgcctt | 6240 |
| ccttgaccct | ggaaggtgcc | actcccactg | tcctttccta | ataaaatgag  | gaaattgcat | 6300 |
| cgcattgtct | gagtaggtgt | cattctattc | tggggggtgg | ggtggggcag  | cacagcaagg | 6360 |
| gggaggattg | ggaagacaat | agcaggcatg | ctggggatgc | ggtgggctct  | atgggtaccc | 6420 |
| aggtgctgaa | gaattgaccc | ggttcctcct | gggccagaaa | gaagcaggca  | catccccttc | 6480 |
| tctgtgacac | accetgteca | cgcccctggt | tcttagttcc | agccccactc  | ataggacact | 6540 |
| catagctcag | gagggctccg | ccttcaatcc | cacccgctaa | agtacttgga  | gcggtctctc | 6600 |
| cctccctcat | cagcccacca | aaccaaacct | agcctccaag | agtgggaaga  | aattaaagca | 6660 |
| agataggcta | ttaagtgcag | agggagagaa | aatgcctcca | acatgtgagg  | aagtaatgag | 6720 |
| agaaatcata | gaatttcttc | cgcttcctcg | ctcactgact | cgctgcgctc  | ggtcgttcgg | 6780 |
| ctgcggcgag | cggtatcagc | tcactcaaag | gcggtaatac | ggttatccac  | agaatcaggg | 6840 |
| gataacgcag | gaaagaacat | gtgagcaaaa | ggccagcaaa | aggccaggaa  | ccgtaaaaag | 6900 |
| gccgcgttgc | tggcgtttt  | ccataggctc | cgcccccctg | acgagcatca  | caaaaatcga | 6960 |
| cgctcaagtc | agaggtggcg | aaacccgaca | ggactataaa | gataccaggc  | gtttccccct | 7020 |
| ggaagctccc | tcgtgcgctc | tcctgttccg | accetgeege | ttaccggata  | cctgtccgcc | 7080 |
| tttctccctt | cgggaagcgt | ggcgctttct | caatgctcac | gctgtaggta  | tctcagttcg | 7140 |
| gtgtaggtcg | ttcgctccaa | gctgggctgt | gtgcacgaac | cccccgttca  | gcccgaccgc | 7200 |
| tgcgccttat | ccggtaacta | tcgtcttgag | tccaacccgg | taagacacga  | cttatcgcca | 7260 |
| ctggcagcag | ccactggtaa | caggattagc | agagcgaggt | atgtaggcgg  | tgctacagag | 7320 |
| ttcttgaagt | ggtggcctaa | ctacggctac | actagaagga | cagtatttgg  | tatctgcgct | 7380 |
| ctgctgaagc | cagttacctt | cggaaaaaga | gttggtagct | cttgatccgg  | caaacaaacc | 7440 |
| accgctggta | gcggtggttt | ttttgtttgc | aagcagcaga | ttacgcgcag  | aaaaaagga  | 7500 |
| tctcaagaag | atcctttgat | cttttctacg | gggtctgacg | ctcagtggaa  | cgaaaactca | 7560 |
| cgttaaggga | ttttggtcat | gagattatca | aaaaggatct | tcacctagat  | ccttttaaat | 7620 |
| taaaaatgaa | gttttaaatc | aatctaaagt | atatatgagt | aaacttggtc  | tgacagttac | 7680 |
| caatgcttaa | tcagtgaggc | acctatctca | gcgatctgtc | tatttcgttc  | atccatagtt | 7740 |
| geetgaetee | 999999999  | ggcgctgagg | tctgcctcgt | gaagaaggtg  | ttgctgactc | 7800 |
| ataccaggee | tgaatcgccc | catcatccag | ccagaaagtg | agggagccac  | ggttgatgag | 7860 |
| agetttgttg | taggtggacc | aqttqqtqat | tttgaacttt | tgctttgcca  | cggaacggtc | 7920 |
| tacattatca | ggaagatgcg | tgatctgatc | cttcaactca | gcaaaagtto  | gatttattca | 7980 |
| acaaaaccac | cgtcccatca | agtcagcgta | atgetetgee | agtgttacaa  | ccaattaacc | 8040 |
|            |            |            |            |             | catatcagga |      |
|            |            |            |            |             | ctcaccgagg |      |
| cagttccata | ggatggcaag | atcctootat | cggtctqcaa | ttccgactcc  | tccaacatca | 8220 |
| atacaaccta | ttaatttccc | ctcqtcaaaa | ataaggttat | caagtgagaa  | atcaccatga | 8280 |
|            |            |            |            |             | J          |      |

```
gtgacgactg aatccggtga gaatggcaaa agcttatgca tttettteca gacttgttea 8340 acaggccage cattacgcte gtcatcaaaa tcactcgcat caaccaaacc gttattcatt 8400 cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt taaaaaggaca attacaaaca 8460 ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa 8520 tcaggatatt cttctaatac ctggaatgct gttttcccgg ggatcgcagt ggtgagtaac 8580 catgcatcat caggagtacg gataaaaatgc ttgatggtcg gaagaggcat aaattccgtc 8640 agccagttta gtctgaccat ctcatctgta acatcattgg caacgctacc tttgccatgt 8700 ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt cgcacctgat 8760 tgcccgacat tatcgcgagc ccatttatac ccatataaat cagcatccat gttggaattt 8820 aatcgcggcc tcgagcaaga cgtttcccgt tgaatatggc tcataacacc ccttgtatta 8880 ctgtttatgt aagcagacag ttttattgtt catgatgata tattttatc ttgtgcaatg 8940 taacaatcaga gattttgaga cacaacgtgg ctttccccc cccccatta ttgaagcatt 9000 ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 9120 atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtc 9169
```

<210> 170
<211> 9792
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 170 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eccetttgge 1080 tettatgeat getatactgt ttttggettg gggeetatae acceegett cettatgeta 1140 taggtgatgg tatagettag cetataggtg tqqqttattq accattattg accactecce 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgac acggactetg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccqtqcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeq agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggecag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

| gcagaagaag  |             |             |            |            |              |      |
|-------------|-------------|-------------|------------|------------|--------------|------|
| gttgcggtgc  | tgttaacggt  | ggagggcagt  | gtagtctgag | cagtactcgt | tgctgccgcg   | 1800 |
| cgcgccacca  | gacataatag  | ctgacagact  | aacagactgt | tcctttccat | gggtctttc    | 1860 |
| tgcagtcacc  | gtcgtcgaca  | tgggcgcccg  | cgccagcgtg | ctgagcggcg | gcgagctgga   | 1920 |
| ccgctgggag  | aagatccgcc  | tgcgccccgg  | cggcaagaag | aagtacaagc | tgaagcacat   | 1980 |
| cgtgtgggcc  | agccgcgagc  | tggagcgctt  | cgccgtgaac | cccggcctgc | tggagaccag   | 2040 |
| cgagggctgc  | cgccagatcc  | tgggccagct  | gcagcccagc | ctgcagaccg | gcagcgagga   | 2100 |
| gctgcgcagc  | ctgtacaaca  | ccgtggccac  | cctgtactgc | gtgcaccagc | gcatcgagat   | 2160 |
|             |             |             | cgaggaggag |            |              |      |
| ggcccagcag  | accaccacca  | acaccggcca  | cagcaaccag | gtgagccaga | actaccccat   | 2280 |
| catacagaac  | atccaqqqcc  | agatggtgca  | ccaggccatc | agcccccgca | ccctgaacgc   | 2340 |
| ctaggtgaag  | ataataaaaa  | agaaggcctt  | cagccccgag | gtgatcccca | tgttcagcgc   | 2400 |
| cctgagcgag  | agcaccaccc  | cccaggacct  | gaacaccatg | ctgaacaccg | tgggcggcca   | 2460 |
| ccadaccacc  | atocaoatoc  | tgaaggagac  | catcaacgag | qaqqccqccq | agtgggaccg   | 2520 |
| catacacccc  | atacacacca  | gccccatcgc  | ccccggccag | atgcgcgagc | cccqcqqcaq   | 2580 |
| cascatege   | dacaccacca  | geaccetgea  | ggagcagatc | gactagatga | ccaacaaccc   | 2640 |
| cccatccc    | ggcaccacca  | tctacaagcg  | ctggatcatc | ctagacctaa | acaagatcut   | 2700 |
| ggggatgtag  | acceccace   | acatectagas | catccgccag | aaccccssaa | agecetteeg   | 2760 |
| gagaatagata | gagggttct   | acaacaccct  | gegegeegag | caddccadcc | aggaggtgaa   | 2820 |
| gactacgtg   | acceptece   | tactaatace  | gaacgccaac | cccactaca  | agaccatcct   | 2880 |
| gaactggatg  | accgagaccc  | ccaccetoga  | ggagatgatg | accacctacc | agacataaa    | 2940 |
| caccacac    | ggccccgccg  | acatactage  | cgaggccatg | accacatga  | ccaacaccc    | 3000 |
| cggccccggc  | atacaaggeee | gcgcgccggc  | caaccagcgc | aagatcatga | agtgcttcaa   | 3060 |
| caccaccacg  | acgcagcgcg  | gcaacccccg  | ctgccgcgcc | ccccccaaga | agagetaeta   | 3120 |
| ctgcggcaag  | gagggccaca  | accacatosa  | ggactgcacc | gaggagga   | ctaataggga   | 3180 |
| gaagtgegge  | thanasasa   | accagacgaa  | ggactgtact | transacana | ccadedgga    | 3240 |
| agatetggee  | tttttaaa    | ggaaggccag  | ggaattttct | aacaactccc | totoagageeaa | 3240 |
| cagececace  | agaagagage  | ctcaggcctg  | gggaagagac | gagatgagtg | tttaacaaaa   | 3360 |
| aggageegat  | agacaaggaa  | bossesses   | tagcttccct | cagattatic | cccggcagcg   | 3430 |
| accectegee  | acaacaaaga  | cagggggcca  | gctgaaggag | tagaaaccaa | acaccggcgc   | 3480 |
| cgacgacacc  | grgerggagg  | agatgaacct  | gcccggccgc | abookoatoo | agacgaccgg   | 3540 |
| cggcatcggc  | ggetteatea  | tactactaca  | gtacgaccag | accongatog | teggeege     | 3600 |
| ccacaaggee  | aceggeaceg  | rgerggeggg  | ccccaccccc | acceceated | agaccataca   | 3660 |
| cetgetgace  | cagacegget  | bassassas   | cttccccatc | ageceeateg | tancagaga    | 3720 |
| cgtgaagetg  | aageeeggea  | eggaeggeee  | caaggtgaag | ageggeeee  | agategagga   | 3700 |
| gaagatcaag  | geeetggtgg  | agatetgeac  | cgagatggag | aaggagggca | agaccagcaa   | 3940 |
| gateggeece  | gagaacccct  | acaacacccc  | cgtgttcgcc | accaagaaga | aggacagcac   | 3040 |
| caagtggcgc  | aagctggtgg  | acttccgcga  | gctgaacaag | cgcacccagg | terestast    | 3900 |
| ggtgcagctg  | ggcatecee   | acceegeegg  | cctgaagcag | aagaagagcg | cgaccgcgcc   | 4020 |
| ggacgrgggc  | gacgcctact  | ccagcgtgcc  | cctggacaag | gaetteegea | agracacege   | 4020 |
| cttcaccatc  | cccagcatca  | acaacgagac  | ccccggcatc | cgctaccagt | acaacgtgct   | 4740 |
| gccccagggc  | rggaagggca  | gccccgccat  | cttccagtgc | ageatgacea | agateetgga   | 4200 |
| gecetteege  | aagcagaacc  | ccgacatcgt  | gatctaccag | cacacggacc | accigiacgi   | 4260 |
| gggcagcgac  | ctggagatcg  | gccagcaccg  | caccaagatc | gaggageege | geeageacet   | 4200 |
| gctgcgctgg  | ggcttcacca  | cccccgacaa  | gaagcaccag | aaggagcccc | cetteetgtg   | 4320 |
| gatgggctac  | gagetgeace  | ccgacaagtg  | gaccgtgcag | cccatcgtgc | Egecegagaa   | 4380 |
| ggacagetgg  | accgtgaacg  | acatccagaa  | gctggtgggc | aagetgaact | gggccagcca   | 4440 |
| gatctacgcc  | ggcatcaagg  | tgcgccagct  | gtgcaagctg | ccgcgcggca | ccaaggccct   | 4500 |
| gaccgaggtg  | grgcccctga  | ccgaggaggc  | cgagctggag | ctggccgaga | accgcgagat   | 4500 |
| cctgaaggag  | cccgtgcacg  | gcgtgtacta  | cgaccccagc | aaggacctga | rcgccgagat   | 4620 |
| ccagaagcag  | ggccagggcc  | agtggaccta  | ccagatctac | caggagccct | tcaagaacct   | 4680 |
| gaagaccggc  | aagtacgccc  | gcatgaaggg  | cgcccacacc | aacgacgtga | agcagctgac   | 4/40 |
| cgaggccgtg  | cagaagatcg  | ccaccgagag  | catcgtgatc | tggggcaaga | ccccaagtt    | 4800 |
| caagctgccc  | atccagaagg  | agacctggga  | ggcctggtgg | accgagtact | ggcaggccac   | 4860 |
| ctggatcccc  | gagtgggagt  | tcgtgaacac  | ccccccctg  | gtgaagetgt | ggtaccaget   | 4920 |
| ggagaaggag  | cccatcatcg  | gcgccgagac  | cttctacgtg | gacggcgccg | ccaaccgcga   | 4980 |

| gaccaagctg | ggcaaggccg | gctacgtgac | cgaccgcggc  | cgccagaagg  | tggtgcccct   | 5040 |
|------------|------------|------------|-------------|-------------|--------------|------|
| gaccgacacc | accaaccaga | agaccgagct | gcaggccatc  | cacctggccc  | tgcaggacag   | 5100 |
|            | gtgaacatcg |            |             |             |              |      |
|            | agcgagagcg |            |             |             |              |      |
| gaaggtgtac | ctggcctggg | tgcccgccca | caagggcatc  | ggcggcaacg  | agcaggtgga   | 5280 |
| cggcctggtg | agcgccggca | tccgcaaggt | gctgttcctg  | gacggcatcg  | acaaggccca   | 5340 |
| ggaggagcac | gagaagtacc | acagcaactg | gcgcgccatg  | gccagcgact  | tcaacctgcc   | 5400 |
| ccccgtggtg | gccaaggaga | tcgtggccag | ctgcgacaag  | tgccagctga  | agggcgaggc   | 5460 |
| catgcacggc | caggtggact | gcagccccgg | catctggcag  | ctggcatgca  | cccacctgga   | 5520 |
| gggcaaggtg | atcctggtgg | ccgtgcacgt | ggccagcggc  | tacatcgagg  | ccgaggtgat   | 5580 |
| cccccccqaq | accggccagg | agaccgccta | cttcctgctg  | aagctggccg  | gccgctggcc   | 5640 |
| cqtqaaqacc | gtgcacaccg | acaacggcag | caacttcacc  | agcaccaccg  | tgaaggccgc   | 5700 |
| ctactaataa | gccggcatca | agcaggagtt | cggcatcccc  | tacaaccccc  | agagccaggg   | 5760 |
| cataatcaaa | agcatgaaca | aggagetgaa | gaagatcatc  | ggccaggtgc  | gcgaccaggc   | 5820 |
| cgagcacctg | aagaccgccg | tacagatage | catattcatc  | cacaacttca  | agcgcaaggg   | 5880 |
| coocatcooc | ggctacagcg | ccaacaaaca | catcatagac  | atcatcgcca  | ccqacatcca   | 5940 |
| gaccaaggag | ctgcagaagc | agatcaccaa | gatccagaac  | ttccacatat  | actaccocoa   | 6000 |
| caccacasc  | cccgtgtgga | agggcccgc  | caagetgetg  | tagaagggg   | agggcgcgt    | 6060 |
| aataataaa  | gacaacagcg | acatcaaggt | gatacccac   | cacaaaacca  | agatcatccg   | 6120 |
| caactacaac | aagcagatgg | ccaccacas  | chacataacc  | agccaccada  | acqaqqacat   | 6180 |
| agaccacage | tggagcaaga | acacatast  | caactaaccc  | accatacaca  | aggggatgg    | 6240 |
|            | cccgccgccg |            |             |             |              |      |
|            | agcagcaaca |            |             |             |              |      |
|            | gaggtgggct |            |             |             |              |      |
| ggaggaggag | gtggacctga | accacttcct | daaddadaad  | aacaacctaa  | agggcctgat   | 6480 |
|            | cgccgccagg |            |             |             |              |      |
|            | cagaactaca |            |             |             |              |      |
|            | ctggtgcccg |            |             |             |              |      |
|            | ctgcaccccg |            |             |             |              |      |
|            | ttcgacagcc |            |             |             |              |      |
| ggagtggcgc | aactgctgag | gatcaattct | tactatacct  | tctagttgcc  | agccatctgt   | 6840 |
|            | tccccgtgc  |            |             |             |              |      |
|            | gaggaaattg |            |             |             |              |      |
|            | cagcacagca |            |             |             |              |      |
| tacaataaa  | tctatgggta | cccacatact | cascasttca  | cccattcct   | cctagaccaa   | 7080 |
|            | gcacatcccc |            |             |             |              |      |
|            | ctcataggac |            |             |             |              |      |
| tanageceta | ggagcggtct | atcoataget | catcaccca   | ccasaccaa   | cctaccetge   | 7260 |
|            | agaaattaaa |            |             |             |              |      |
|            | aggaagtaat |            |             |             |              |      |
|            | ctcggtcgtt |            |             |             |              |      |
|            |            |            |             |             |              |      |
|            | cacagaatca |            |             |             |              |      |
|            | gaaccgtaaa |            |             |             |              |      |
| ctgacgagca | tcacaaaaat | cgacgcccaa | gccagaggcg  | gegaaaceeg  | acaygactat   | 7620 |
|            | ggcgtttccc |            |             |             |              |      |
|            | atacctgtcc |            |             |             |              |      |
|            | gtatctcagt |            |             |             |              |      |
|            | tcagcccgac |            |             |             |              |      |
|            | cgacttatcg |            |             |             |              |      |
|            | cggtgctaca |            |             |             |              |      |
| ggacagtatt | tggtatctgc | accounted  | ayccayccac  | tttttt      | tacaaaaaa    | 8100 |
| getettgate | cggcaaacaa | accaccgerg | gragraggrag | gatottttgct | acconnectate | 8150 |
| agactacgcg | cagaaaaaaa | yyarcecaag | aayatttta   | garcultuct  | hasssaga     | 0700 |
| acgeceageg | gaacgaaaac | nathannah  | gyacttegge  | atrastata   | agtatatata   | 9220 |
| ccccaccta  | gateetttta | aaccaaaaac | gaagttttaa  | accaatttaa  | agracacacg   | 0200 |

```
agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 8340
gtctatttcg ttcatccata gttgcctgac tccggggggg gggggcgctg aggtctgcct 8400
cgtgaagaag gtgttgctga ctcataccag gcctgaatcg ccccatcatc cagccagaaa 8460
gtgagggagc cacggttgat gagagctttg ttgtaggtgg accagttggt gattttgaac 8520
ttitgctttg ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg atccttcaac 8580
tcagcaaaag ttcgatttat tcaacaaagc cgccgtcccg tcaagtcagc gtaatgctct 8640
gccagtgtta caaccaatta accaattctg attagaaaaa ctcatcgagc atcaaatgaa 8700
actgcaattt attcatatca ggattatcaa taccatattt ttgaaaaagc cgtttctgta 8760
atgaaggaga aaactcaccg aggcagttcc ataggatggc aagatcctgg tatcggtctg 8820
cgattccgac tcgtccaaca tcaatacaac ctattaattt cccctcgtca aaaataaggt 8880
tatcaagtga gaaatcacca tgagtgacga ctgaatccgg tgagaatggc aaaagcttat 8940
gcatttettt ccagaettgt tcaacaggee agecattacg etegtcatca aaatcacteg 9000
catcaaccaa accettatte atteeteatt geeceteage gagacgaaat acgegatege 9060
tgttaaaagg acaattacaa acaggaatcg aatgcaaccg gcgcaggaac actgccagcg 9120
catcaacaat attttcacct gaatcaggat attcttctaa tacctggaat gctgttttcc 9180
cggggatcgc agtggtgagt aaccatgcat catcaggagt acggataaaa tgcttgatgg 9240
toggaagagg cataaattoo gtoagooagt ttagtotgac catotoatot gtaacatoat 9300
tggcaacgct acctttgcca tgtttcagaa acaactctgg cgcatcgggc ttcccataca 9360
atcgatagat tgtcgcacct gattgcccga cattatcgcg agcccattta tacccatata 9420
aatcagcatc catgttggaa tttaatcgcg gcctcgagca agacgtttcc cgttgaatat 9480
ggctcataac accccttgta ttactgttta tgtaagcaga cagttttatt gttcatgatg 9540
atatatttt atcttgtgca atgtaacatc agagattttg agacacaacg tggctttccc 9600
cccccccca ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg 9660
aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 9720
ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga 9780
qqccctttcg tc
```

<211> 9167 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 171 tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tecaacatta cegecatgit gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660

ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960 tagaagacac cgggaccgat ccagcctcc cggccggaa cggtgcattg gaacgcggat 1020

<210> 171

|            |            |            |            |            | cccctttggc  |      |
|------------|------------|------------|------------|------------|-------------|------|
|            |            |            |            |            | ccttatgcta  |      |
|            |            |            |            |            | accactcccc  |      |
|            |            |            |            |            | caactatctc  |      |
|            |            |            |            |            | tatttttaca  |      |
|            |            |            |            |            | cccccgtgcc  |      |
|            |            |            |            |            | gtgttccgga  |      |
| catgggctct | tctccggtag | cggcggagct | tccacatccg | agccctggtc | ccatgcctcc  | 1500 |
|            |            |            |            |            | acttaggcac  |      |
|            |            |            |            |            | gtatgtgtct  |      |
| gaaaatgagc | gtggagattg | ggctcgcacg | gctgacgcag | atggaagact | taaggcagcg  | 1680 |
|            |            |            |            |            | ggtaactccc  |      |
| gttgcggtgc | tgttaacggt | ggagggcagt | gtagtctgag | cagtactcgt | tgctgccgcg  | 1800 |
| cgcgccacca | gacataatag | ctgacagact | aacagactgt | tcctttccat | gggtcttttc  | 1860 |
| tgcagtcacc | gtcgtcgaca | tggcccgcgc | cagcgtgctg | agcggcggcg | agctggaccg  | 1920 |
| ctgggagaag | atccgcctgc | gccccggcgg | caagaagaag | tacaagctga | agcacatcgt  | 1980 |
| gtgggccagc | cgcgagctgg | agcgcttcgc | cgtgaacccc | ggcctgctgg | agaccagcga  | 2040 |
| gggctgccgc | cagatcctgg | gccagctgca | gcccagcctg | cagaccggca | gcgaggagct  | 2100 |
| gcgcagcctg | tacaacaccg | tggccaccct | gtactgcgtg | caccagcgca | tcgagatcaa  | 2160 |
| ggacaccaag | gaggccctgg | acaagatcga | ggaggagcag | aacaagagca | agaagaaggc  | 2220 |
|            |            |            |            |            | accccatcgt  |      |
| gcagaacatc | cagggccaga | tggtgcacca | ggccatcagc | ccccgcaccc | tgaacgcctg  | 2340 |
|            |            |            |            |            | tcagcgccct  |      |
|            |            |            |            |            | gcggccacca  |      |
|            |            |            |            |            | gggaccgcgt  |      |
|            |            |            |            |            | gcggcagcga  |      |
|            |            |            |            |            | acaaccccc   |      |
|            |            |            |            |            | agatcgtgcg  |      |
|            |            |            |            |            | ccttccgcga  |      |
|            |            |            |            |            | aggtgaagaa  |      |
|            |            |            |            |            | ccatcctgaa  |      |
|            |            |            |            |            | gcgtgggcgg  |      |
|            |            |            |            |            | acagcgccac. |      |
|            |            |            |            |            | gcttcaactg  |      |
|            |            |            |            |            | gctgctggaa  |      |
|            |            |            |            |            | atagggaaga  |      |
|            |            |            |            |            | gagccaacag  |      |
|            |            |            |            |            | cagaagcagg  |      |
|            |            |            |            |            | ggcagcgacc  |      |
|            |            |            |            |            | ccggcgccga  |      |
|            |            |            |            |            | tgatcggcgg  |      |
|            |            |            |            |            | tctgcggcca  |      |
|            |            |            |            |            | gccgcaacct  |      |
|            |            |            |            |            | ccgtgcccgt  |      |
|            |            |            |            |            | ccgaggagaa  |      |
|            |            |            |            |            | tcagcaagat  |      |
|            |            |            |            |            | acagcaccaa  |      |
|            |            |            |            |            | tctgggaggt  |      |
|            |            |            |            |            | ccgtgctgga  |      |
|            |            |            |            |            | acaccgcctt  |      |
|            |            |            |            |            | acgtgctgcc  |      |
|            |            |            |            |            | tcctggagcc  |      |
|            |            |            |            |            | tgtacgtggg  |      |
|            |            |            |            |            | agcacctgct  |      |
|            |            |            |            |            | tcctgtggat  |      |
| 2-2-22226  |            |            |            |            |             |      |

gggctacgag ctgcaccccg acaagtggac cgtgcagccc atcgtgctgc ccgagaagga 4380 cagetggace gtgaaegaca tecagaaget ggtgggcaag etgaaetggg ceagecagat 4440 ctacgccggc atcaaggtgc gccagctgtg caagctgctg cgcggcacca aggccctgac 4500 cgaggtggtg cccctgaccg aggaggccga gctggagctg gccgagaacc gcgagatcct 4560 gaaggagccc gtgcacggcg tgtactacga ccccagcaag gacctgatcg ccgagatcca 4620 gaagcagggc cagggccagt ggacctacca gatctaccag gagcccttca agaacctgaa 4680 gaccggcaag tacgcccgca tgaagggcgc ccacaccaac gacgtgaagc agctgaccga 4740 ggccgtgcag aagatcgcca ccgagagcat cgtgatctgg ggcaagaccc ccaagttcaa 4800 gctgcccatc cagaaggaga cctgggaggc ctggtggacc gagtactggc aggccacctg 4860 gateceegag tgggagtteg tgaacaeece eeeeetggtg aagetgtggt accagetgga 4920 gaaggageee ateateggeg eegagaeett etaegtggae ggegeegeea aeegegagae 4980 caagctgggc aaggccggct acgtgaccga ccgcggccgc cagaaggtgg tgcccctgac 5040 egacaceace aaccagaaga eegagetgea ggecateeac etggeeetge aggacagegg 5100 cetggaggtg aacategtga eegacageca gtacgeeetg ggcatcatec aggeecagee 5160 cgacaagagc gagagcgagc tggtgagcca gatcatcgag cagctgatca agaaggagaa 5220 ggtgtacctg gcctgggtgc ccgcccacaa gggcatcggc ggcaacgagc aggtggacgg 5280 cctggtgagc gccggcatcc gcaaggtgct gttcctggac ggcatcgaca aggcccagga 5340 ggagcacgag aagtaccaca gcaactggcg cgccatggcc agcgacttca acctgccccc 5400 cgtggtggcc aaggagatcg tggccagctg cgacaagtgc cagctgaagg gcgaggccat 5460 gcacggccag gtggactgca gccccggcat ctggcagctg gcatgcaccc acctggaggg 5520 caaggtgate etggtggeeg tgeacgtgge cageggetae ategaggeeg aggtgatece 5580 egecgagace ggecaggaga eegectaett eetgetgaag etggeeggee getggeeegt 5640 gaagaccgtg cacaccgaca acggcagcaa cttcaccagc accaccgtga aggccgcctg 5700 ctggtgggcc ggcatcaagc aggagttcgg catcccctac aacccccaga gccagggcgt 5760 gatcgagagc atgaacaagg agctgaagaa gatcatcggc caggtgcgcg accaggccga 5820 gcacctgaag accgccgtgc agatggccgt gttcatccac aacttcaagc gcaagggcgg 5880 categgegge tacagegeeg gegagegeat egtggacate ategecaceg acatecagae 5940 caaggagetg cagaagcaga tcaccaagat ccagaacttc cgcgtgtact accgcgacag 6000 ccgcgacccc gtgtggaagg gccccgccaa gctgctgtgg aagggcgagg gcgccgtggt 6060 gatccaggac aacagcgaca tcaaggtggt gccccgccgc aaggccaaga tcatccgcga 6120 ctacggcaag cagatggccg gcgacgactg cgtggccagc cgccaggacg aggactagga 6180 attettgetg tgeettetag ttgeeageea tetgttgttt geeceteeee egtgeettee 6240 ttgaccctgg aaggtgccac tcccactgtc ctttcctaat aaaatgagga aattgcatcg 6300 cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagca cagcaagggg 6360 gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat gggtacccag 6420 gtgctgaaga attgacccgg ttcctcctgg gccagaaaga agcaggcaca tccccttctc 6480 tgtgacacac cotgtocacg cocotggtto ttagttocag cocoactcat aggacactca 6540 tageteagga gggeteegee tteaateeea eeegetaaag taettggage ggteteteee 6600 teceteatea geccaceaaa ecaaacetag eetecaagag tgggaagaaa ttaaagcaag 6660 ataggctatt aagtgcagag ggagagaaaa tgcctccaac atgtgaggaa gtaatgagag 6720 aaatcataga atttetteeg etteeteget caetgaeteg etgegetegg tegttegget 6780 geggegageg gtateagete acteaaagge ggtaataegg ttatecaeag aateagggga 6840 taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 6900 egegttgetg gegtttttee ataggeteeg eeceetgae gageateaca aaaategaeg 6960 ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 7020 aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 7080 tetecetteg ggaagegtgg egetttetea atgeteacge tgtaggtate teagtteggt 7140 gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 7200 cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact 7260 ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 7320 cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct 7380 gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 7440 cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 7500 tcaagaagat cetttgatet tttctacggg gtctgacget cagtggaacg aaaactcacg 7560 ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 7620

```
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 7680
atgettaate agtgaggeae etateteage gatetgteta tttegtteat ecatagttge 7740
ctgactcegg gggggggggg cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat 7800
accaggeetg aategeeeca teatecagee agaaagtgag ggageeaegg ttgatgagag 7860
ctttgttgta ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg 7920
cgttgtcggg aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac 7980
aaagccgccg tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa 8040
ttctgattag aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt 8100
atcaatacca tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca 8160
gttccatagg atggcaagat cctggtatcg gtctgcgatt ccgactcgtc caacatcaat 8220
acaacctatt aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt 8280
gacgactgaa tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac 8340
aggecageca ttacgetegt catcaaaate actegeatea accaaacegt tatteatteg 8400
tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg 8460
aatcgaatgc aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc 8520
aggatattet tetaataeet ggaatgetgt ttteeegggg ategeagtgg tgagtaacea 8580
tgcatcatca ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag 8640
ccagtttagt ctgaccatct catctgtaac atcattggca acgetacctt tgccatgttt 8700
cagaaacaac tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg 8760
cccgacatta tcgcgagccc atttataccc atataaatca gcatccatgt tggaatttaa 8820
tegeggeete gageaagaeg ttteeegttg aatatggete ataacaeece ttgtattact 8880
gtttatgtaa gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta 8940
acatcagaga ttttgagaca caacgtggct ttccccccc ccccattatt gaagcattta 9000
tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat 9060
aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat 9120
                                                                  9167
catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtc
```

```
<210> 172
<211> 9783
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
<400> 172
togogogttt oggtgatgac ggtgaaaaco totgacacat gcagotocog gagacggtca 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ceegectggc tgacegeca acgaeeceeg cecattgacg teaataatga egtatgttee 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
```

| teceegtgee  |              | taataaaa                    | atatagacto   | tataddcaca   | ccctttaac  | 1080 |
|-------------|--------------|-----------------------------|--------------|--------------|------------|------|
| teceegtgee  | aagagtgatg   | taagtactgc                  | ccacagacce   | acceptate    | ccttatgcta | 1140 |
| tcttatgcat  | getataetge   |                             | teachtatta   | accettgee    | accactccc  | 1200 |
| taggtgatgg  | tatagettag   | cctataggtg                  | tyggttattg   | atcattaceg   | caactatete | 1260 |
| tattggtgac  | gatactttcc   | actactaatc                  | cataacatgg   | accaseteta   | tatttttaca | 1320 |
| tattggctat  | atgccaatac   | tetgteette                  | agagactgac   | acggactetg   | cacccataca | 1380 |
| ggatggggtc  | ccatttatta   | tttacaaatt                  | cacatataca   | totaggetag   | ghathacaga | 1440 |
| cgcagttttt  | attaaacata   | gcgtgggatc                  | tccacgcgaa   | tctcgggtac   | gegeeeegga | 1500 |
| catgggctct  | teteeggtag   | cggcggagct                  | tecacateeg   | agccctggtc   | anthagasa  | 1560 |
| agcggctcat  | ggtcgctcgg   | cagctccttg                  | cccctaacag   | tggaggccag   | accuaggeac | 1620 |
| agcacaatgc  | ccaccaccac   | cagtgtgccg                  | cacaaggccg   | tggcggtagg   | beargages  | 1600 |
| gaaaatgagc  | gtggagattg   | ggctcgcacg                  | gctgacgcag   | atggaagact   | caaggeageg | 1000 |
| gcagaagaag  | atgcaggcag   | ctgagttgtt                  | gtattctgat   | aagagtcaga   | ggtaactccc | 1/40 |
| gttgcggtgc  | tgttaacggt   | ggagggcagt                  | gtagtctgag   | cagtactcgt   | tgctgccgcg | 1800 |
| cgcgccacca  | gacataatag   | ctgacagact                  | aacagactgt   | tcctttccat   | adarcttrc  | 1860 |
| t.gcagtcacc | gtcgtcgaca   | tggcccgcgc                  | cagcgtgctg   | agcggcggcg   | agctggaccg | 1920 |
| ctgggagaag  | atccgcctgc   | gccccggcgg                  | caagaagaag   | tacaagctga   | agcacatcgt | 1980 |
| gtgggccagc  | cgcgagctgg   | agcgcttcgc                  | cgtgaacccc   | ggcctgctgg   | agaccagcga | 2040 |
| gggctgccgc  | cagatcctgg   | gccagctgca                  | gcccagcctg   | cagaccggca   | gcgaggagct | 2100 |
| gegeageetg  | tacaacaccg   | tggccaccct                  | gtactgcgtg   | caccagcgca   | tcgagatcaa | 2160 |
| ggacaccaag  | gaggccctgg   | acaagatcga                  | ggaggagcag   | aacaagagca   | agaagaaggc | 2220 |
| ccagcaggcc  | gccgccgaca   | ccggccacag                  | caaccaggtg   | agccagaact   | accccatcgt | 2280 |
| gcagaacatc  | cagggccaga   | tggtgcacca                  | ggccatcagc   | ccccgcaccc   | tgaacgcctg | 2340 |
| ggtgaaggtg  | gtggaggaga   | aggccttcag                  | ccccgaggtg   | atccccatgt   | tcagcgccct | 2400 |
| gagcgagggc  | gccacccccc   | aggacctgaa                  | caccatgctg   | aacaccgtgg   | gcggccacca | 2460 |
| ggccgccatg  | cagatgctga   | aggagaccat                  | caacgaggag   | gccgccgagt   | gggaccgcgt | 2520 |
| gcaccccqtg  | cacgccggcc   | ccatcgcccc                  | cggccagatg   | cgcgagcccc   | gcggcagcga | 2580 |
| categeegge  | accaccagca   | ccctgcagga                  | gcagatcggc   | tggatgacca   | acaacccccc | 2640 |
| catccccqtq  | ggcgagatct   | acaagcgctg                  | gatcatcctg   | ggcctgaaca   | agatcgtgcg | 2700 |
| catqtacaqc  | cccaccagca   | tcctggacat                  | ccgccagggc   | cccaaggagc   | ccttccgcga | 2760 |
| ctacqtqqac  | cgcttctaca   | agaccctgcg                  | cgccgagcag   | gccagccagg   | aggtgaagaa | 2820 |
| ctggatgacc  | gagaccctgc   | tggtgcagaa                  | cgccaacccc   | gactgcaaga   | ccatcctgaa | 2880 |
| gaccetagge  | cccgccgcca   | ccctggagga                  | gatgatgacc   | gcctgccagg   | gcgtgggcgg | 2940 |
| ccccqqccac  | aaggcccgcg   | tgctggccga                  | ggccatgagc   | caggtgacca   | acagcgccac | 3000 |
| catcatgatg  | cagcgcggca   | acttccgcaa                  | ccagcgcaag   | atcgtgaagt   | gcttcaactg | 3060 |
| caacaaggag  | ggccacaccg   | cccgcaactg                  | ccgcgccccc   | cgcaagaagg   | gctgctggaa | 3120 |
| atacaacaaa  | gagggccacc   | agatgaagga                  | ctgcaccgag   | cgacaggcta   | atagggaaga | 3180 |
| tetageette  | ccacaaqqqa   | aggccaggga                  | attttcttca   | gagcagacca   | gagccaacag | 3240 |
| cccaccaga   | agagagette   | aggtttgggg                  | aaqaqacaac   | aactccctct   | cagaagcagg | 3300 |
| agccgataga  | caaggaactg   | tatcctttag                  | cttccctcag   | atcactcttt   | ggcagcgacc | 3360 |
| cctcgtcaca  | ataaagatag   | gagaccaact                  | gaaggaggcc   | cttctagaca   | ccggcgccga | 3420 |
| caacaccata  | ctggaggaga   | tgaacctgcc                  | cagecactag   | aagcccaaga   | tgatcggcgg | 3480 |
| categgegg   | ttcatcaago   | tagaccagta                  | cqaccagato   | ctgatcgaga   | tctgcggcca | 3540 |
| caecagogge  | ogcaccatac   | taataaacco                  | cacccccqtq   | aacatcatcg   | gccgcaacct | 3600 |
| actascosa   | atcaactaca   | ccctgaactt                  | cccatcago    | cccatcgaga   | ccgtgcccgt | 3660 |
| geegaceeag  | cccaacataa   | acaacccaa                   | gataaaacaa   | tggcccctga   | ccqaqqaqaa | 3720 |
| gaageegaag  | ctaggaaagg   | , tetgesecca                | gatggagaag   | gagggcaaga   | tcagcaagat | 3780 |
| gattaaggtt  | accectace    | acaccccat                   | : attcaccato | aagaagaagg   | acagcaccaa | 3840 |
| etgeceegag  | adcccccaca   | tecaraaart                  | gaacaagcgc   | acccaggact   | tctgggaggt | 3900 |
| gryycycaag  | aboccccac    | . ccaccaacct                | : gaagragaar | aagagcgtga   | ccatactaaa | 3960 |
| geagetggge  | accidentact  | , cogoeggeet                | . ddacaacdac | : ttccgcaagt | acaccacctt | 4020 |
| cgrgggcgac  | geetaette    | , gegegeeee                 | . Jacaayyac  | taccagtaca   | acatactaca | 4080 |
| caccatcccc  | . agcatcaaca | . acgagaccc                 | . coaghgeage | atgaccaaga   | tectagage  | 4140 |
| ccayggctgg  | aagygcagcc   | , cogocacoco<br>, acatemtos | ctaccantac   | atggaccacc   | tatacataa  | 4200 |
| curegeaag   | g dagaaccccc | , acattytyai                | r caadatcoa  | gagetgegee   | agcacctgct | 4260 |
| cayogaccco  | gagaccggc    | . agcaccycac                | a descesas   | , andeceeece | tectatagat | 4320 |
| gcgctgggg   | : ttcaccacco | . ccyacaayaa                | . youccayaag | g gagcccccct |            |      |

| agactacgag  | chacacaca  | acaagtagac | catacaacce | atcgtgctgc | ccqaqaaqqa  | 4380 |
|-------------|------------|------------|------------|------------|-------------|------|
| cagctggacc  | gtgaacgaca | tccagaagct | ggtgggcaag | ctgaactggg | ccaqccaqat  | 4440 |
| ctaccegace  | atcaacatac | accaactata | caagetgetg | cgcggcacca | aggccctgac  | 4500 |
| caacataata  | ccctaacca  | accaaccaa  | actagaacta | gccgagaacc | acqaqatcct  | 4560 |
| caaggeggeg  | apacagaaga | totactacoa | cccadcaad  | gacctgatcg | ccgagatcca  | 4620 |
| gaaggageee  | gegeacggeg | agacctacga | gatchaccag | gagcccttca | agaacctgaa  | 4680 |
| gaageaggge  | tagggccage | tanagagag  | gacccaccag | gageeeeea  | adchascas   | 4740 |
| gaceggeaag  | Lacyceegea | cgaagggcgc | ccacaccaac | gacgtgaagc | ccaacttcaa  | 4800 |
| ggccgtgcag  | aagaccgcca | ccgagagcat | cycyacccyy | ggcaagaccc | accaggggaa  | 4000 |
| getgeecate  | cagaaggaga | cctgggagge | ccggcggace | gagtactggc | aggeeacetg  | 4020 |
| gatccccgag  | tgggagttcg | tgaacacccc | ccccctggtg | aagctgtggt | accagetgga  | 4920 |
| gaaggagccc  | atcatcggcg | ccgagacctt | ccacgtggac | ggcgccgcca | accegegagae | 4900 |
| caagctgggc  | aaggccggct | acgtgaccga | ccgcggccgc | cagaaggtgg | tgeecetgae  | 5040 |
| cgacaccacc  | aaccagaaga | ccgagctgca | ggccatccac | ctggccctgc | aggacagegg  | 5100 |
| cctggaggtg  | aacatcgtga | ccgacagcca | gtacgccctg | ggcatcatcc | aggcccagcc  | 2700 |
| cgacaagagc  | gagagcgagc | tggtgagcca | gatcatcgag | cagctgatca | agaaggagaa  | 5220 |
| ggtgtacctg  | gcctgggtgc | ccgcccacaa | gggcatcggc | ggcaacgagc | aggtggacgg  | 5280 |
| cctggtgagc  | gccggcatcc | gcaaggtgct | gttcctggac | ggcatcgaca | aggcccagga  | 5340 |
| ggagcacgag  | aagtaccaca | gcaactggcg | cgccatggcc | agcgacttca | acctgccccc  | 5400 |
| cgtggtggcc  | aaggagatcg | tggccagctg | cgacaagtgc | cagctgaagg | gcgaggccat  | 5460 |
| gcacggccag  | gtggactgca | gccccggcat | ctggcagctg | gcatgcaccc | acctggaggg  | 5520 |
| caaggtgatc  | ctggtggccg | tgcacgtggc | cagcggctac | atcgaggccg | aggtgatccc  | 5580 |
| cgccgagacc  | ggccaggaga | ccgcctactt | cctgctgaag | ctggccggcc | gctggcccgt  | 5640 |
| gaagaccgtg  | cacaccgaca | acggcagcaa | cttcaccagc | accaccgtga | aggccgcctg  | 5700 |
| ctggtgggcc  | ggcatcaagc | aggagttcgg | catcccctac | aacccccaga | gccagggcgt  | 5760 |
| gatcgagagc  | atgaacaagg | agctgaagaa | gatcatcggc | caggtgcgcg | accaggccga  | 5820 |
| gcacctgaag  | accgccgtgc | agatggccgt | gttcatccac | aacttcaagc | gcaagggcgg  | 5880 |
| catcggcggc  | tacagcgccg | gcgagcgcat | cgtggacatc | atcgccaccg | acatccagac  | 5940 |
| caaggagetg  | cagaagcaga | tcaccaagat | ccagaacttc | cgcgtgtact | accgcgacag  | 6000 |
| ccqcgacccc  | gtgtggaagg | gccccgccaa | gctgctgtgg | aagggcgagg | gcgccgtggt  | 6060 |
| gatccaggac  | aacagcgaca | tcaaggtggt | gccccgccgc | aaggccaaga | tcatccgcga  | 6120 |
| ctacggcaag  | cagatggccg | gcgacgactg | cgtggccagc | cgccaggacg | aggacatggg  | 6180 |
| cggcaagtgg  | agcaagagca | gcgtgatcgg | ctggcccgcc | gtgcgcgagc | gcatgcgccg  | 6240 |
| cgccgagccc  | gccgccgacg | gcgtgggcgc | cgtgagccgc | gacctggaga | agcacggcgc  | 6300 |
| catcaccage  | agcaacaccg | ccgccaacaa | cgccgcctgc | gcctggctgg | aggcccagga  | 6360 |
| ggaggaggag  | gtgggcttcc | ccgtgacccc | ccaggtgccc | ctgcgcccca | tgacctacaa  | 6420 |
| ggccgccgtg  | gacctgagcc | acttcctgaa | ggagaagggc | ggcctggagg | gcctgatcca  | 6480 |
| cagccagcgc  | cgccaggaca | tcctggacct | gtggatctac | cacacccagg | gctacttccc  | 6540 |
| cgactggcag  | aactacaccc | ccggccccgg | cgtgcgctac | cccctgacct | tcggctggtg  | 6600 |
| ctacaagetg  | gtgcccgtgg | agcccgacaa | ggtggaggag | gccaacaagg | gcgagaacac  | 6660 |
| cagcctgctg  | caccccgtga | gcctgcacgg | catggacgac | cccgagcgcg | aggtgctgga  | 6720 |
| gtggcgcttc  | gacagccgcc | tggccttcca | ccacgtggcc | cgcgagctgc | accccgagta  | 6780 |
| cttcaagaac  | tgctgaggat | caattcttgc | tgtgccttct | agttgccagc | catctgttgt  | 6840 |
| ttqcccctcc  | cccgtgcctt | ccttgaccct | ggaaggtgcc | actcccactg | tcctttccta  | 6900 |
| ataaaatgag  | gaaattgcat | cgcattgtct | gagtaggtgt | cattctattc | tggggggtgg  | 6960 |
| gatagagacaa | cacagcaagg | gggaggattg | ggaagacaat | agcaggcatg | ctggggatgc  | 7020 |
| gatagactct  | atgggtaccc | aggtgctgaa | gaattgaccc | ggttcctcct | gggccagaaa  | 7080 |
| gaagcaggca  | catccccttc | tctqtqacac | accctqtcca | cgcccctggt | tcttagttcc  | 7140 |
| agccccactc  | ataggacact | catageteag | qaqqqctccq | ccttcaatcc | cacccgctaa  | 7200 |
| agtacttoga  | gcggtctctc | cctccctcat | cageccaeca | aaccaaacct | agcctccaaq  | 7260 |
| agtgggaaga  | aattaaagca | agataggeta | ttaaqtqcaq | agggagagaa | aatgcctcca  | 7320 |
| acatotoaco  | aagtaatgag | agaaatcata | gaatttette | cgcttcctcg | ctcactgact  | 7380 |
| cactacacta  | gatcattcaa | ctgcggcgag | cggtatcagg | tcactcaaag | gcggtaatac  | 7440 |
| gattatecae  | agaatcaggg | gataacgcag | qaaaqaacat | gtgagcaaaa | ggccagcaaa  | 7500 |
| addccaddaa  | ccgtaaaaaa | accacattac | tagcatttt  | ccataggete | cgccccctq   | 7560 |
| acgaggatga  | caaaaatcoa | cqctcaaqtc | agaggtggcg | aaacccgaca | ggactataaa  | 7620 |
|             |            | J          |            |            |             |      |

```
gataccagge gtttccccct ggaagetece tegtgegete teetgtteeg accetgeege 7680
ttaccggata cctgtccgcc tttctccctt cgggaagegt ggcgctttct caatgctcac 7740
getqtaqqta teteagtteg gtgtaggteg ttegetecaa getgggetgt gtgcaegaac 7800
ccccqttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg 7860
taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 7920
atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga 7980
cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct 8040
cttgatccgg caaacaaacc accgctggta gcggtggtgt ttgcaagcag cagattacgc 8100
gcaqaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgctcagt 8160
ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct 8220
agatcetttt aaattaaaaa tgaagtttta aatcaateta aagtatatat gagtaaaett 8280
ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc 8340
gttcatccat agttgcctga ctccgggggg ggggggcgct gaggtctgcc tcgtgaagaa 8400
ggtgttgctg actcatacca ggcctgaatc gccccatcat ccagccagaa agtgagggag 8460
ccacggttga tgagagettt gttgtaggtg gaccagttgg tgattttgaa ettttgettt 8520
gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct gatccttcaa ctcagcaaaa 8580
qttcqattta ttcaacaaag ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt 8640
acaaccaatt aaccaattct gattagaaaa actcatcgag catcaaatga aactgcaatt 8700
tattcatatc aggattatca ataccatatt tttgaaaaag ccgtttctgt aatgaaggag 8760
aaaactcacc gaggcagttc cataggatgg caagatcctg gtatcggtct gcgattccga 8820
ctcgtccaac atcaatacaa cctattaatt tcccctcgtc aaaaataagg ttatcaagtg 8880
aqaaatcacc atqaqtgacq actgaatccg gtgagaatgg caaaagctta tgcatttctt 8940
tccagacttg ttcaacaggc cagccattac gctcgtcatc aaaatcactc gcatcaacca 9000
aaccqttatt cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaag 9060
gacaattada aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa 9120
tattttcacc tgaatcagga tattcttcta atacctggaa tgctgttttc ccggggatcg 9180
caqtqqtqaq taaccatqca tcatcaqqaq tacqqataaa atqcttqatg gtcggaagag 9240
qcataaattc cgtcagccag tttagtctga ccatctcatc tgtaacatca ttggcaacgc 9300
tacctttgcc atgtttcaga aacaactctg gcgcatcggg cttcccatac aatcgataga 9360
ttgtcgcacc tgattgcccg acattatcgc gagcccattt atacccatat aaatcagcat 9420
ccatgttgga atttaatcgc ggcctcgagc aagacgtttc ccgttgaata tggctcataa 9480
caccccttgt attactgttt atgtaagcag acagttttat tgttcatgat gatatatttt 9540
tatcttgtgc aatgtaacat cagagatttt gagacacaac gtggctttcc ccccccccc 9600
attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 9660
agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgtct 9720
aagaaaccat tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc 9780
```

```
<210> 173
<211> 9782
<212> DNA
<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 173
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
```

ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 eccgcetgge tgacegeeca acgaceceeg cecattgaeg teaataatga egtatgttee 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtetec accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataaqcaq 900 ageteghtta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaayacac egggacegat ecageeteeg eggeegggaa eggtgeattg gaaegeggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eeeetttgge 1.080 tettatgeat getatactgt ttttggettg gggeetatac accecegett cettatgeta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteoggtag eggeggaget tecacateeg ageeetggte ecatgeetee 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 egegecacea gacataatag etgacagaet aacagaetgt teetttecat gggtetttte 1860 tgcagtcacc gtcgtcgaca tggcaaagtg gagcaagagc agcgtgatcg gctggcccgc 1920 cgtgcgcgag cgcatgcgcc gcgccgagcc cgccgccgac ggcgtgggcg ccgtgagccg 1980 cgacctggag aagcacggcg ccatcaccag cagcaacacc gccgccaaca acgccqcctq 2040 egectggetg gaggeceagg aggaggagga ggtgggette eeegtgaeee eeeaggtgee 2100 cetgegeece atgacetaea aggeegeegt ggacetgage caetteetga aggagaaggg 2160 eggeetggag ggeetgatee acageeageg eegeeaggae atcetggaee tgtggateta 2220 ccacacccag ggctacttcc ccgactggca gaactacacc cccggccccg gcgtgcgcta 2280 cccctgacc ttcggctggt gctacaagct ggtgcccgtg gagcccgaca aggtggagga 2340 ggccaacaag ggcgagaaca ccagcctgct gcaccccgtg agcctgcacg gcatggacga 2400 ccccgagcgc gaggtgctgg agtggcgctt cgacagccgc ctggccttcc accacgtggc 2460 cegegagetg caccegagt actteaagaa etgeatggee egegeeageg tgetgagegg 2520 eggegagetg gacegetggg agaagateeg eetgegeeee ggeggeaaga agaagtacaa 2580 getgaageae ategtgtggg ceageegega getggagege thegeegtga acceeggeet 2640 getggagace agegaggget geegeeagat cetgggeeag etgeageeea geetgeagae 2700 cggcagcgag gagctgcgca gcctgtacaa caccgtggcc accctgtact gcgtgcacca 2760 gcgcatcgag atcaaggaca ccaaggaggc cctggacaag atcgaggagg agcagaacaa 2820 gagcaagaag aaggcccagc aggccgccgc cgacaccggc cacagcaacc aggtgagcca 2880 gaactacccc atcgtgcaga acatccaggg ccagatggtg caccaggcca tcagcccccg 2940 caccctgaac gcctgggtga aggtggtgga ggagaaggcc ttcagccccg aggtgatccc 3000 catgttcagc gccctgagcg agggcgccac cccccaggac ctgaacacca tgctgaacac 3060 cgtgggcggc caccaggccg ccatgcagat gctgaaggag accatcaacg aggaggccgc 3120 egagtgggac egegtgeace eegtgeacge eggeeecate geeceeggee agatgegega 3180 geceegegge agegacateg eeggeaceae cageaceetg eaggageaga teggetggat 3240 gaccaacaac ccccccatcc ccgtgggcga gatctacaag cgctggatca tcctgggcct 3300 gaacaagate gtgegeatgt acageeecae cageateetg gacateegee agggeeecaa 3360 ggagcccttc cgcgactacg tggaccgctt ctacaagacc ctgcgcgccg agcaggccag 3420 ccaggaggtg aagaactgga tgaccgagac cctgctggtg cagaacgcca accccgactg 3480 caagaccatc ctgaaggccc tgggccccgc cgccaccctg gaggagatga tgaccgcctg 3540 ccagggcgtg ggcggccccg gccacaaggc ccgcgtgctg gccgaggcca tgagccaggt 3600 gaccaacage gecaecatea tgatgeageg eggeaactte egeaaceage geaagategt 3660

| gaagtgcttc | aactgcggca | aggagggcca | caccgcccgc | aactgccgcg | cccccgcaa                | 3720    |
|------------|------------|------------|------------|------------|--------------------------|---------|
|            |            |            |            |            | ccgagcgaca               |         |
|            |            |            |            |            | cttcagagca               |         |
|            |            |            |            |            | acaacaactc               |         |
|            |            |            |            |            | ctcagatcac               |         |
|            |            |            |            |            | aggcccttct               |         |
|            |            |            |            |            | gctggaagcc               |         |
|            |            |            |            |            | agatcctgat               |         |
|            |            |            |            |            | ccgtgaacat               |         |
| categgeege | aacctgctga | cccagatcgg | ctgcaccctg | aacttcccca | tcagccccat               | 4260    |
| cgagaccgtg | cccgtgaagc | tgaagcccgg | catggacggc | cccaaggtga | agcagtggcc               | 4320    |
| cctgaccgag | gagaagatca | aggccctggt | ggagatctgc | accgagatgg | agaaggaggg               | 4380    |
|            |            |            |            |            | ccatcaagaa               |         |
|            |            |            |            |            | agcgcaccca               |         |
| ggacttctgg | gaggtgcagc | tgggcatccc | ccaccccgcc | ggcctgaagc | agaagaagag               | 4560    |
| cgtgaccgtg | ctggacgtgg | gcgacgccta | cttcagcgtg | cccctggaca | aggacttccg               | 4620    |
|            |            |            |            |            | tccgctacca               |         |
|            |            |            |            |            | gcagcatgac               |         |
|            |            |            |            |            | agtacatgga               |         |
|            |            |            |            |            | tcgaggagct               |         |
|            |            |            |            |            | agaaggagcc               |         |
|            |            |            |            |            | agcccatcgt               |         |
|            |            |            |            |            | gcaagctgaa               |         |
|            |            |            |            |            | tgctgcgcgg               |         |
|            |            |            |            |            | agctggccga               |         |
|            |            |            |            |            | gcaaggacct               |         |
|            |            |            |            |            | accaggagcc               |         |
|            |            |            |            |            | ccaacgacgt               |         |
| gaagcagctg | accgaggccg | tgcagaagat | cgccaccgag | agcatcgtga | tctggggcaa               | 5400    |
|            |            |            |            |            | ggaccgagta               |         |
|            |            |            |            |            | tggtgaagct               |         |
|            |            |            |            |            | tggacggcgc               |         |
|            |            |            |            |            | gccgccagaa               |         |
|            |            |            |            |            | tccacctggc               |         |
|            |            |            |            |            | ccctgggcat               |         |
|            |            |            |            |            | tcgagcagct               |         |
|            |            |            |            |            | tcggcggcaa               |         |
|            |            |            |            |            | tggacggcat               |         |
|            |            |            |            |            | tggccagcga               |         |
|            |            |            |            |            | agtgccagct               |         |
|            |            |            |            |            | agctggcatg               |         |
|            |            |            |            |            | gctacatcga               |         |
|            |            |            |            |            | tgaagetgge               |         |
|            |            |            |            |            | ccagcaccac<br>cctacaaccc |         |
|            |            |            |            |            | tcggccaggt               |         |
|            |            |            |            |            | tccacaactt               |         |
|            |            |            |            |            | acatcatcgc               |         |
|            |            |            |            |            | acttccgcgt               |         |
|            |            |            |            |            | tgtggaaggg               |         |
|            |            |            |            |            | gccgcaaggg               |         |
|            |            |            |            |            | ccagccgcca               |         |
|            |            |            |            |            | tgtttgccc                |         |
|            |            |            |            |            | ctaataaaat               |         |
|            |            |            |            |            | tggggtgggg               |         |
| 5-55-4009  |            |            | - 5        |            | - 2222433                | J J J J |

|            | -          |            |             |              | 101/0      |      |
|------------|------------|------------|-------------|--------------|------------|------|
|            |            |            |             |              | tgcggtgggc |      |
| tctatgggta | cccaggtgct | gaagaattga | cccggttcct  | cctgggccag   | aaagaagcag | 7080 |
|            |            |            |             |              | tccagcccca |      |
|            |            |            |             |              | taaagtactt |      |
| ggagcggtct | ctccctccct | catcagccca | ccaaaccaaa  | cctagcctcc   | aagagtggga | 7260 |
| agaaattaaa | gcaagatagg | ctattaagtg | cagagggaga  | gaaaatgcct   | ccaacatgtg | 7320 |
| aggaagtaat | gagagaaatc | atagaatttc | ttccgcttcc  | tegeteactg   | actcgctgcg | 7380 |
| ctcggtcgtt | cggctgcggc | gagcggtatc | agctcactca  | aaggcggtaa   | tacggttatc | 7440 |
|            |            |            |             |              | aaaaggccag |      |
|            |            |            |             |              | ctgacgagca |      |
|            |            |            |             |              | aaagatacca |      |
|            |            |            |             |              | cgcttaccgg |      |
| atacctgtcc | gcctttctcc | cttcgggaag | cgtggcgctt  | tctcaatgct   | cacgctgtag | 7740 |
| gtatctcagt | toggtgtagg | tegttegete | caagctgggc  | tgtgtgcacg   | aacccccgt  | 7800 |
| tcagcccgac | cactacacct | tatccggtaa | ctatcgtctt  | gagtccaacc   | cggtaagaca | 7860 |
| cgacttatcg | ccactggcag | cagccactgg | taacaggatt  | agcagagcga   | ggtatgtagg | 7920 |
| caatactaca | gagttcttga | agtggtggcc | taactacggc  | tacactagaa   | ggacagtatt | 7980 |
| tagtatetae | gctctgctga | agccagttac | cttcggaaaa  | agagttggta   | gctcttgatc | 8040 |
|            |            |            |             |              | agattacgcg |      |
|            |            |            |             |              | acgctcagtg |      |
| даасдаааас | tcacgttaag | ggattttggt | catgagatta  | tcaaaaaqqa   | tcttcaccta | 8220 |
| gatcctttta | aattaaaaat | gaagttttaa | atcaatctaa  | agtatatatg   | agtaaacttg | 8280 |
|            |            |            |             |              | gtctatttcg |      |
|            |            |            |             |              | cgtgaagaag |      |
|            |            |            |             |              | gtgagggagc |      |
|            |            |            |             |              | ttttgctttg |      |
|            |            |            |             |              | tcagcaaaag |      |
| ttccatttat | tcaacaaaac | caccatacca | tcaagtcagc  | gtaatgetet   | gccagtgtta | 8640 |
| caaccaatta | accaattctq | attagaaaaa | ctcatcgage  | atcaaatgaa   | actgcaattt | 8700 |
|            |            |            |             |              | atgaaggaga |      |
|            |            |            |             |              | cgattccgac |      |
|            |            |            |             |              | tatcaagtga |      |
|            |            |            |             |              | gcatttcttt |      |
|            |            |            |             |              | catcaaccaa |      |
|            |            |            |             |              | tgttaaaagg |      |
|            |            |            |             |              | catcaacaat |      |
|            |            |            |             |              | cggggatcgc |      |
|            |            |            |             |              | tcggaagagg |      |
|            |            |            |             |              | tggcaacgct |      |
|            |            |            |             |              | atcgatagat |      |
| tateacacat | gettgccga  | cattatogg  | agcccattta  | tacccatata   | aatcagcatc | 9420 |
| catattages | tttaateee  | acctaceges | ageceatetee | cattaeatat   | ggctcataac | 9480 |
| acceptions | thactothe  | totaaccaca | cachtttatt  | attesteste   | atatatttt  | 9540 |
| atchtatas  | atataacata | agagattta  | agacacaacc  | tagetttee    | cccccccca  | 9600 |
| ttattanaa  | abthatasa  | attattatat | catgacaacg  | tacatattto   | aatgtattta | 9660 |
|            |            |            |             |              | ctgacgtcta |      |
|            |            |            |             |              | ggccctttcg |      |
| tc         | accaccacga | Caccaaccaa | Lududacagg  | og ca coacga | 22000000   | 9782 |
|            |            |            |             |              |            | 2.04 |

<210> 174

<211> 9788

<212> DNA

<213> Artificial Sequence

<220>
<223> plasmid pVR1012x/s containing HIV genes

<400> 174 tegegegttt eggtgatgac ggtgaaaace tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcgggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaacegte agategeetg gagaegeeat ceaegetgtt ttgaeeteea 960 tagaagacac cgggaccgat ccagcetecg cggccgggaa cggtgcattg gaacgcggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaea eecetttgge 1080 tottatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tetgteette agagactgae aeggactetg tatttttaca 1320 ggatggggte ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500 ageggeteat ggtegetegg eageteettg etectaacag tggaggeeag aettaggeae 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggtcttttc 1860 tgcagtcacc gtcgtcgaca tgggcgcccg cgccagcatc ctgcgcggcg gcaagctgga 1920 cqcctqqqaq cqcatcaaqc tgaagcccgg cggcaagaag cactacatga tgaagcacct 1980 ggtgtgggcc agccgcgagc tggagcgctt cgccctggac cccggcctgc tggagaccag 2040 cgagggctgc aagcagatca tgaagcagct gcagcccgcc ctgcagaccg gcaccaagga 2100 gctgatcagc ctgcacaaca ccgtggccac cctgtactgc gtgcacgaga agatcgacgt 2160 gcgcgacacc aaggaggccc tggacaagat caaggaggag cagaacaaga gccagcagaa 2220 gacccagcag gccgaggccg ccgacaaggg caaggtgagc cagaactacc ccatcgtgca 2280 gaacctgcag ggccagatgg tgcaccagcc catcagcgcc cgcaccetga acgcctgggt 2340 gaaggtggtg gaggagaagg cetteageee egaggtgate eccatgttea eegeeetgag 2400 cgagggcgcc accccccagg acctgaacac catgctgaac accgtgggcg gccaccaggc 2460 cgccatgcag atgctgaagg acaccatcaa cgaggaggcc gccgagtggg accgcctgca 2520 ccccgtgcac gccggccccg tggcccccgg ccagatgcgc gagccccgcg gcagcgacat 2580 cgccggcacc accagcaccc tgcaggagca gatcacctgg atgaccaaca acccccccgt 2640 gcccgtgggc gacatctaca agcgctggat catcctgggc ctgaacaaga tcgtgcgcat 2700 gtacageece gtgageatee tggacateaa geagggeece aaggageeet teegegaeta 2760 cgtggaccgc ttcttcaaga ccctgcgcgc cgagcaggcc acccaggacg tgaagaactg 2820 gatgaccgac accetgetgg tgeagaacge caacceegac tgeaagacca teetgegege 2880 cctgggccc ggcgccagcc tggaggagat gatgaccgcc tgccagggcg tgggcgccc 2940 cggccacaag gcccgcgtgc tggccgaggc catgagcaag gtgaacaaca ccaacatcat 3000

| antaanaaaa | aggaagtaga | 2000000022 | aaaaaaaata | aaataattaa | actgcggcaa | 3060 |
|------------|------------|------------|------------|------------|------------|------|
|            |            |            |            |            |            |      |
|            |            |            |            |            | ggaagtgcgg |      |
|            |            |            |            |            | aagatctggc |      |
|            |            |            |            |            | acagccccac |      |
|            |            |            |            |            | caggagccga |      |
|            |            |            |            |            | gacccctcgt |      |
| cacaataaag | atagggggcc | agctgaagga | ggcccttcta | gacaccggcg | ccgacgacac | 3420 |
| cgtgctggag | gagatgaacc | tgcccggccg | ctggaagccc | aagatgatcg | gcggcatcgg | 3480 |
|            |            |            |            |            | gccacaaggc |      |
|            |            |            |            |            | acctgctgac |      |
|            |            |            |            |            | ccgtgaagct |      |
|            |            |            |            |            | agaagatcaa |      |
|            |            |            |            |            | agatcggccc |      |
|            |            |            |            |            | ccaagtggcg |      |
|            |            |            |            |            | aggtgcagct |      |
|            |            |            |            |            | tggacgtggg |      |
|            |            |            |            |            | ccttcaccat |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | tgccccaggg |      |
| ctggaagggc | ageceegeea | terrecagig | caycatgacc | aagateetgg | agcccttccg | 4140 |
|            |            |            |            |            | tgggcagcga |      |
| cctggagatc | ggccagcacc | gcaccaagat | cgaggagctg | cgccagcacc | tgctgcgctg | 4260 |
| gggcttcacc | acccccgaca | agaagcacca | gaaggagccc | cccttcctgt | ggatgggcta | 4320 |
|            |            |            |            |            | aggacagctg |      |
|            |            |            |            |            | agatctacgc |      |
| cggcatcaag | gtgcgccagc | tgtgcaagct | gctgcgcggc | accaaggccc | tgaccgaggt | 4500 |
| ggtgcccctg | accgaggagg | ccgagctgga | gctggccgag | aaccgcgaga | tcctgaagga | 4560 |
| gcccgtgcac | ggcgtgtact | acgaccccag | caaggacctg | atcgccgaga | tccagaagca | 4620 |
|            |            |            |            |            | tgaagaccgg |      |
|            |            |            |            |            | ccgaggccgt |      |
|            |            |            |            |            | tcaagctgcc |      |
|            |            |            |            |            | cctggatccc |      |
|            |            |            |            |            | tggagaagga |      |
|            |            |            |            |            | agaccaagct |      |
|            |            |            |            |            | tgaccgacac |      |
|            |            |            |            |            | gcggcctgga |      |
|            |            |            |            |            | agcccgacaa |      |
|            |            |            |            |            | agaaggtgta |      |
|            |            |            |            |            | acggcctggt |      |
|            |            |            |            |            | aggaggagca |      |
|            |            |            |            |            | ccccgtggt  |      |
|            |            |            |            |            |            |      |
|            |            |            |            |            | ccatgcacgg |      |
|            |            |            |            |            | agggcaaggt |      |
|            |            |            |            |            | tccccgccga |      |
|            |            |            |            |            | ccgtaaagac |      |
|            |            |            |            |            | cctgctggtg |      |
| ggccggcatc | aagcaggagt | teggeatece | ctacaacccc | cagagccagg | gcgtgatcga | 5760 |
|            |            |            |            |            | ccgagcacct |      |
|            |            |            |            |            | gcggcatcgg |      |
| cggctacagc | gccggcgagc | gcatcgtgga | catcatcgcc | accgacatcc | agaccaagga | 5940 |
|            |            |            |            |            | acagccgcga |      |
|            |            |            |            |            | tggtgatcca |      |
|            |            |            |            |            | gcgactacgg |      |
|            |            |            |            |            | tgggtaataa |      |
|            |            |            |            |            | gccgcgcccc |      |
| cgcagccgaa | ggagtqqqtc | cagetteaca | ggatagtgac | aaatacggcg | ctctgacatc | 6300 |
|            |            | -          |            |            | •          |      |

cagcagtacc cccgctaaca acgccgactg tgcatggctc gaagcccagc aggaagaaga 6360 ggaggtegga ttecetgtga ggeeteaggt teeteteaga eetatgaett acaaggeegt 6420 ggtcgacttg tcttttttcc ttgaagagaa gggaggcttg gaggggctga tttacagcaa 6480 gaagcgccag gatattttgg atttgtgggt ctacaacaca cagggttact tccccgattg 6540 gcagaactac accccggcc caggcgtgcg gtttcctctt acctttggtt ggtgctttaa 6600 getegtgeet gtegateeta gagaagtgga ggaageaaae acaggegaga acaacagtet 6660 getgeatect atgageette atggeatgga ggatteteat egggaagtge tteagtggaa 6720 attegatage tigetegeta gaagacacat ggetegggag tigeateetg agtattataa 6780 ggattqttga acacgtggcg aattettget gtgcctteta gttgccagec atetgttgtt 6840 tgcccctccc ccgtgccttc cttgaccctg gaaggtgcca ctcccactgt cctttcctaa 6900 gtggggcagc acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgcg 7020 gtgggctcta tgggtaccca ggtgctgaag aattgacccg gttcctcctg ggccagaaag 7080 aagcaggcac atccccttct ctgtgacaca ccctgtccac gcccctggtt cttagttcca 7140 geoceactea taggacaete atageteagg agggeteege etteaateee acceqetaaa 7200 gtacttggag cggtctctcc ctccctcatc agcccaccaa accaaaccta gcctccaaga 7260 gtgggaagaa attaaagcaa gataggctat taagtgcaga gggagagaaa atgcctccaa 7320 catgtgagga agtaatgaga gaaatcatag aatttettee getteetege teactgaete 7380 getgegeteg gtegttegge tgeggegage ggtateaget caetcaaagg eggtaatacg 7440 gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaa gccagcaaaa 7500 ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 7560 cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 7620 ataccagged tttecccctg gaageteect egtgegetet cetgtteega eeetgeeget 7680 tacoggatac obstoccook thouseout gggaagogtg gogothtobo aatgeboacq 7740 ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 7800 ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 7860 aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 7920 tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac 7980 agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 8040 ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 8100 tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 8160 tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 8220 cacctagate ettttaaatt aaaaatgaag ttttaaatea atetaaagta tatatgagta 8280 aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 8340 atttcgttca tccatagttg cctgactccg ggggggggg gcgctgaggt ctgcctcgtg 8400 aagaaggtgt tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga 8460 gggagccacg gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt 8520 getttgecae ggaacggtet gegttgtegg gaagatgegt gatetgatee tteaacteag 8580 caaaagttcg atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca 8640 gtgttacaac caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg 8700 caatttattc atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga 8760 aggagaaaac tcaccgaggc agttccatag gatggcaaga tcctggtatc ggtctgcgat 8820 tecgaetegt ceaacateaa tacaacetat taattteece tegteaaaaa taaggttate 8880 aagtgagaaa tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat 8940 ttctttccag acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc 9000 aaccaaaccg ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt 9060 aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc 9120 aacaatattt tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg 9180 gatcgcagtg gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg 9240 aagaggcata aatteegtea gecagtttag tetgaceate teatetgtaa cateattgge 9300 aacgctacct ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg 9360 atagattgtc gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc 9420 agcatccatg ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct 9480 cataacaccc cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat 9540 atttttatct tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc 9600

cccccattat tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg 9660 tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga 9720 cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc 9780 ctttcgtc 9788

<210> 175 <211> 9407 <212> DNA <213> Artificial Sequence <220> <223> plasmid pVR1012x/s containing HIV genes <400> 175 tegegegtht eggtgatgae ggtgaaaacc tetgacacat geageteeeg gagaeggtea 60 cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420 cocqcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720 catcaatggg cgtggatagc ggtttgactc acgggggattt ccaagtctcc accccattga 7.80 cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840 ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900 agetegttta gtgaaccgte agategeetg gagaegeeat ceaegetgtt ttgaceteea 960 tagaagacac egggacegat ecageeteeg eggeegggaa eggtgeattg gaaegeggat 1020 teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eeeetttgge 1080 tottatgcat gotatactgt ttttggottg gggootatac accoccoctt cottatgcta 1140 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440 catgggetet teteeggtag eggeggaget tecacateeg agecetggte ceatgeetee 1500 ageggeteat ggtegetegg cageteettg etectaacag tggaggeeag acttaggeac 1560 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620 qaaaatgagc qtqqaqattq qqctcqcacg gctgacgcag atqqaaqact taaggcagcg 1680 qcaqaagaaq atqcaqqcaq ctqaqttqtt gtattctqat aaqaqtcaqa qqtaactccc 1740 gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800 cgcgccacca qacataataq ctqacagact aacagactqt teetttecat qqqtetttte 1860 tgcagtcacc gtcgtcgaca tggcccgcgc cagcgtgctg agcggcggcg agctggaccg 1920 ctgggagaag atccgcctgc gccccggcgg caagaagaag tacaagctga agcacatcgt 1980 gtgggccage cgcgagctgg agcgcttcgc cgtgaacccc ggcctgctgg agaccagcga 2040 gggctgccgc cagatcctgg gccagctgca gcccagcctg cagaccggca gcgaggagct 2100 gegeageetg tacaacaceg tggccaceet gtactgegtg caccagegea tegagateaa 2160 ggacaccaag gaggccctgg acaagatcga ggaggagcag aacaagagca agaagaaggc 2220 ccagcaggee geegeegaca ceggeeacag caaccaggtg agecagaact accccategt 2280

gcagaacate cagggecaga tggtgcacca ggccateage ceeegcacee tgaacgeetg 2340

|            | gtggaggaga |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
|            | gccacccccc |            |            |            |            |      |
|            | cagatgctga |            |            |            |            |      |
|            | cacgccggcc |            |            |            |            |      |
|            | accaccagca |            |            |            |            |      |
| catccccgtg | ggcgagatct | acaagcgctg | gatcatcctg | ggcctgaaca | agatcgtgcg | 2700 |
| catgtacagc | cccaccagca | tcctggacat | ccgccagggc | cccaaggagc | ccttccgcga | 2760 |
| ctacgtggac | cgcttctaca | agaccctgcg | cgccgagcag | gccagccagg | aggtgaagaa | 2820 |
| ctggatgacc | gagaccctgc | tggtgcagaa | cgccaacccc | gactgcaaga | ccatcctgaa | 2880 |
| ggccctgggc | cccgccgcca | ccctggagga | gatgatgacc | gcctgccagg | gcgtgggcgg | 2940 |
| ccccggccac | aaggcccgcg | tgctggccga | ggccatgagc | caggtgacca | acagcgccac | 3000 |
| catcatgatg | cagcgcggca | acttccgcaa | ccagcgcaag | atcgtgaagt | gcttcaactg | 3060 |
| cggcaaggag | ggccacaccg | cccgcaactg | ccgcgccccc | cgcaagaagg | gctgctggaa | 3120 |
| gtgcggcaag | gagggccacc | agatgaagga | ctgcaccgag | cgacaggcta | ataacgccgc | 3180 |
|            | ctggaggccc |            |            |            |            |      |
|            | cccatgacct |            |            |            |            |      |
| gggcggcctg | gagggcctga | tccacagcca | gcgccgccag | gacatcctgg | acctgtggat | 3360 |
|            | cagggctact |            |            |            |            |      |
|            | accttcggct |            |            |            |            |      |
|            | aagggcgaga |            |            |            |            |      |
|            | cgcgaggtgc |            |            |            |            |      |
|            | ctgcaccccg |            |            |            |            |      |
|            | ctggaggaga |            |            |            |            |      |
|            | ttcatcaagg |            |            |            |            |      |
|            | ggcaccgtgc |            |            |            |            |      |
|            | ateggetgea |            |            |            |            |      |
|            | cccggcatgg |            |            |            |            |      |
|            | ctggtggaga |            |            |            |            |      |
|            | aacccctaca |            |            |            |            |      |
|            | ctggtggact |            |            |            |            |      |
|            | atccccacc  |            |            |            |            |      |
|            | gcctacttca |            |            |            |            |      |
|            | agcatcaaca |            |            |            |            |      |
|            | aagggcagcc |            |            |            |            |      |
|            | cagaaccccg |            |            |            |            |      |
|            | gagateggee |            |            |            |            |      |
|            | ttcaccaccc |            |            |            |            |      |
|            | ctgcaccccg |            |            |            |            |      |
|            | gtgaacgaca |            |            |            |            |      |
|            | atcaaggtgc |            |            |            |            |      |
| caaaataata | ccctgaccg  | aggaggcga  | actagaacta | gccgagaacc | gcgagatcct | 4800 |
| gaaggagccc | gtgcacggcg | totactacga | ccccagcaag | gacctgatcg | ccgagatcca | 4860 |
|            | cagggccagt |            |            |            |            |      |
|            | tacgcccgca |            |            |            |            |      |
|            | aagatcgcca |            |            |            |            |      |
|            | cagaaggaga |            |            |            |            |      |
|            | tgggagttcg |            |            |            |            |      |
|            | atcatcggcg |            |            |            |            |      |
|            | aaggccggct |            |            |            |            |      |
|            | aaccagaaga |            |            |            |            |      |
|            | aacatcgtga |            |            |            |            |      |
|            | gagagcgagc |            |            |            |            |      |
|            | gcctgggtgc |            |            |            |            |      |
|            | gccggcatcc |            |            |            |            |      |
|            | aagtaccaca |            |            |            |            |      |
| ggagcacgag | aagtattata | geaactygeg | ogceatggee | agegactica | acceptace  | 2040 |

| cgtggtggcc | aaggagatcg  | tggccagctg | cgacaagtgc  | cagctgaagg | gcgaggccat  | 5700 |
|------------|-------------|------------|-------------|------------|-------------|------|
| gcacggccag | gtggactgca  | gccccggcat | ctggcagctg  | gcatgcaccc | acctggaggg  | 5760 |
| caaggtgatc | ctggtggccg  | tgcacgtggc | cagcggctac  | atcgaggccg | aggtgatccc  | 5820 |
| cqccqaqacc | ggccaggaga  | ccgcctactt | cctgctgaag  | ctggccggcc | gctggcccgt  | 5880 |
|            |             |            | cttcaccagc  |            |             |      |
| ctaataaacc | ggcatcaagc  | aggagttcgg | catecectae  | aacccccaga | gccagggcgt  | 6000 |
| gatcgagagc | atgaacaagg  | agctgaagaa | gatcatcggc  | caggtgcgcg | accaggccga  | 6060 |
| gcacctgaag | accoccotoc  | agatggccgt | gttcatccac  | aacttcaagc | gcaagggcgg  | 6120 |
| categgegge | tacagegeeg  | gcgagcgcat | cgtggacatc  | atcgccaccg | acatccagac  | 6180 |
| caaggagctg | cagaagcaga  | tcaccaagat | ccagaacttc  | cgcgtgtact | accgcgacag  | 6240 |
|            |             |            | gctgctgtgg  |            |             |      |
|            |             |            | gccccgccgc  |            |             |      |
| ctacggcaag | cagatogoco  | gcgacgactg | cgtggccagc  | caccaaaaca | aggactagga  | 6420 |
| attettacta | taccttctag  | ttgccagcca | tctgttgttt  | accetteec  | catacettee  | 6480 |
|            |             |            | ctttcctaat  |            |             |      |
|            |             |            | gggggtgggg  |            |             |      |
| gaggattggg | aagacaatag  | caggcatgct | ggggatgcgg  | toggetetat | gagtacccag  | 6660 |
| gugguccggg | attoaccco   | thechected | gccagaaaga  | agcaggcaca | tecettete   | 6720 |
| tatazazaa  | cctatccaca  | ccctaattc  | ttagttccag  | cccactcat  | aggacactca  | 6780 |
|            |             |            | cccgctaaag  |            |             |      |
|            |             |            | cctccaagag  |            |             |      |
| cteeteatea | geccaccaaa  | ggagagaaaa | tgcctccaac  | atotoaggaa | ataataaaa   | 6960 |
| acaggerace | adytycagag  | attactact  | cactgactcg  | ctacacteaa | teatteaget  | 7020 |
| aaaccacaga | attitettetg | actorcages | ggtaatacgg  | ttatecaeae | aatcaccaca  | 7020 |
| geggegageg | gtateagete  | acccaaagge | ggtaatacgg  | accadasacc | attananggga | 7140 |
| taacgcagga | aagaacatgt  | gagcaaaagg | ccagcaaaaag | gccaggaacc | graaaaaggc  | 7200 |
| cgcgttgctg | gegtttttee  | acaggereeg | ccccctgac   | baggaccaca | thosasstas  | 7260 |
| ctcaagtcag | aggtggcgaa  | accegacagg | actataaaga  | aggastage  | tatagagatt  | 7220 |
| aagctccctc | grgegerere  | etgiteegae | cctgccgctt  | tatagatata | tarattarat  | 7320 |
|            |             |            | atgeteacge  |            |             |      |
|            |             |            | gcacgaaccc  |            |             |      |
| cgccttatcc | ggtaactate  | gccccgagce | caacccggta  | agacacgact | abassast    | 7500 |
| ggcagcagcc | actggtaaca  | ggattagcag | agcgaggtat  | graggeggrg | tatagagit   | 7500 |
|            |             |            | tagaaggaca  |            |             |      |
|            |             |            | tggtagctct  |            |             |      |
|            |             |            | gcagcagatt  |            |             |      |
|            |             |            | gtctgacgct  |            |             |      |
| ttaagggatt | ttggtcatga  | gattatcaaa | aaggatcttc  | acctagatec | ccccaaacca  | 7860 |
|            |             |            | atatgagtaa  |            |             |      |
|            |             |            | gatctgtcta  |            |             |      |
| ctgactccgg | aaaaaaaaaa  | cgctgaggtc | tgcctcgtga  | agaaggtgtt | getgaeteat  | 8040 |
| accaggcctg | aatcgcccca  | tcatccagcc | agaaagtgag  | ggagccacgg | ttgatgagag  | 8100 |
| ctttgttgta | ggtggaccag  | ttggtgattt | tgaacttttg  | ctttgccacg | gaacggtctg  | 8790 |
| cgttgtcggg | aagatgcgtg  | atctgatcct | tcaactcagc  | aaaagttcga | tttattcaac  | 8220 |
| aaagccgccg | tecegteaag  | tcagcgtaat | gctctgccag  | tgttacaacc | aattaaccaa  | 8280 |
| ttctgattag | aaaaactcat  | cgagcatcaa | atgaaactgc  | aatttattca | tatcaggatt  | 8340 |
| atcaatacca | tatttttgaa  | aaagccgttt | ctgtaatgaa  | ggagaaaact | caccgaggca  | 8400 |
| gttccatagg | atggcaagat  | cctggtatcg | gtctgcgatt  | ccgactcgtc | caacatcaat  | 8460 |
|            |             |            | aaggttatca  |            |             |      |
|            |             |            | cttatgcatt  |            |             |      |
|            |             |            | actcgcatca  |            |             |      |
|            |             |            | atcgctgtta  |            |             |      |
|            |             |            | cagcgcatca  |            |             |      |
| aggatattct | tctaatacct  | ggaatgctgt | tttcccgggg  | atcgcagtgg | tgagtaacca  | 8820 |
| tgcatcatca | ggagtacgga  | taaaatgctt | gatggtcgga  | agaggcataa | attccgtcag  | 8880 |
| ccagtttagt | ctgaccatct  | catctgtaac | atcattggca  | acgctacctt | tgccatgttt  | 8940 |
|            |             |            |             |            |             |      |

cagaaacaac tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg 9000

```
cccgacatta tcgcgagccc atttataccc atataaatca gcatccatgt tggaatttaa 9060
tegeggeete gageaagacg tttecegttg aatatggete ataacacccc ttgtattact 9120
gtttatgtaa gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta 9180
acatcagaga ttttgagaca caacgtggct ttcccccccc ccccattatt gaagcattta 9240
tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat 9300
aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat 9360
catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtc
<210> 17,6
<211> 9785
<212> DNA
<213> Artificial Sequence
<220>
<223> plasmid pVR1012x/s containing HIV genes
tegegegttt eggtgatgae ggtgaaaace tetgacacat geageteeeg gagaeggtea 60
cagettgtet gtaageggat geegggagea gacaageeeg teagggegeg teagegggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tecaacatta cegecatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agetegitta gigaacegie agaicgeeig gagaegeeai eeaegeigit tigaeeteea 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020
teccegtgee aagagtgaeg taagtaeege etatagaete tataggeaca eccetttgge 1080
tettatgeat getatactgt ttttggettg gggeetatac acceegett cettatgeta 1140
taggtgatgg tatagettag cetataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tetgteette agagaetgae acggaetetg tatttttaea 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggetct teteeggtag eggeggaget tecacateeg agecetggte ceatgetee 1500
ageggeteat ggtegetegg eageteettg etectaaeag tggaggeeag aettaggeae 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga qqtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
egegecacea gacataatag etgacagaet aacagaetgt teetttecat gggtetttte 1860
tgcagtcacc gtcgtcgaca tgggcgcccg cgccagcgtg ctgagcggcg gcaagttcga 1920
cgcctgggag aagatccgcc tgcgccccgg cggcaagaag aagtaccgca tgaagcacct 1980
gatctgggcc agccgcgagc tggaccgctt cgccctgaac cccagcctgc tggagaccgc 2040
cgacggctgc cagcagatca tggagaagct gcagcccgcc ctgaagaccg gcaccgagga 2100
```

| gatcaagagc  | ctgttcaaca               | ccgtggccac | cctgtactgc | gtgcaccagc | gcatcgacgt | 2160  |
|-------------|--------------------------|------------|------------|------------|------------|-------|
| gaaggacacc  | aaggaggccc               | tggacaagat | cgaggagatc | aagaacaaga | gcaagcagaa | 2220  |
| gacccagcag  | gccgccgccg               | acaccggcaa | cagcagcaac | gtgagccaga | actaccccat | 2280  |
| cgtgcagaac  | gcccagggcc               | agatgatcca | ccagagcctg | agcccccqca | ccctgaacgc | 2340  |
| ctgggtgaag  | gtgatcgagg               | agaaggcctt | cagccccqag | qtqatcccca | tattcagcac | 2400  |
| cctgagcgag  | ggcgccaccc               | cccaggacct | gaacatgatg | ctgaacatcg | tagacaacca | 2460  |
| ccaggccgcc  | atgcagatgc               | tgaaggacac | catcaacgag | gaggccgccg | agtgggaccg | 252Ó  |
| cctgcacccc  | gtgcacgccg               | gccccatccc | ccccqqccaq | atgcgcgagc | cccacaacaa | 2580  |
| cgacatcgcc  | ggcaccacca               | gcacccccca | ggagcagatc | gactagataa | ccadcaaccc | 2640  |
| cccatcccc   | gtgggcgaca               | tctacaagcg | ctggatcatc | ctgggcctga | acaagatcot | 2700  |
| gcgcatgtac  | agccccgtga               | gcatcctgga | catcaagcag | ggcccaagg  | agcccttccg | 2760  |
| cgactacgtg  | gaccgcttct               | tcaagacctt | ccacaccaaa | caggccaccc | aggacgtgaa | 2820  |
| gaactggatg  | accgacaccc               | tactaataca | gaacgccaac | cccgactgca | aggccatcct | 2880  |
| acacacceta  | ggccccggcg               | ccaccctgga | ggagatgatg | accocctocc | aggacataga | 2940  |
| caaccccaac  | cacaaggccc               | acatactaac | cgaggccatg | agccaggtgc | adcacaccaa | 3000  |
| catcatgatg  | cagcgcggca               | acttcaaggg | ccadaaccc  | atcaagtgct | tcaactccaa | 3060  |
| caaggaggg   | cacctggccc               | gcaactgcg  | caccctacac | aagaagggct | actageage  | 3120  |
| caacaaaaa   | ggccaccaga               | tgaaggactg | cacccacca  | caggetaata | googgaageg | 3120  |
| ggccttccca  | caagggaagg               | ccanquaatt | ttcttcagag | caggecaata | gggaagaccc | 3740  |
| caccacacaca | gagcttcagg               | tttaaaaaa  | anacaacaac | teestetese | aaaaaaaaaa | 2200  |
| coatagacaa  | ggaactgtat               | cctttagctt | ccctcacatc | actetttese | aagcaggagc | 3360  |
| cgacagacaa  | aanatanna                | acceattae  | gaagggatt  | accectegge | agegaeeeet | 3360  |
| caccatacta  | aagatagggg               | acctacacaa | ggaggccccc | ccagacaccg | gegeegaega | 3420  |
| caccaccttc  | gaggagatga<br>atcaaggtgg | acceptecgg | ccgctggaag | abogagatga | ceggeggeat | 3480  |
| agecateac   | accatactag               | tagaacaaa  | ccagateeta | accyagaccc | geggeeacaa | 3540  |
| ggccaccggc  | accgtgctgg               | tassattaa  | ccccgcgaac | attateggee | geaacetget | 3600  |
| gatteaagatt | ggctgcaccc               | aggaggagga | gaaggeeee  | accgagaccg | tgcccgtgaa | 3660  |
| geegaageee  | ggcatggacg               | gccccaaggt | gaagcagtgg | cccctgaccg | aggagaagat | 3/20  |
| caaggccccg  | gtggagatct               | geacegagat | ggagaaggag | ggcaagatea | gcaagatcgg | 3/80  |
| accessage   | ccctacaaca               | agazatas   | egecateaag | aagaaggaca | gcaccaagtg | 3840  |
| gegeaagetg  | gtggacttcc               | gegagetgaa | caagegeaee | caggacttet | gggaggtgca | 3900  |
| gergggeare  | cccaccccg                | ceggeetgaa | gcagaagaag | agegtgaceg | tgctggacgt | 3960  |
| gggegaegee  | tacttcagcg               | cgcccccgga | caaggacttc | cgcaagtaca | ccgccttcac | 4020  |
| catecccage  | atcaacaacg               | agacccccgg | cateegetae | cagtacaacg | tgctgccca  | 4080. |
| gggccggaag  | ggcagccccg               | ccatetteca | gradeard   | accaagatee | tggagecett | 4140  |
| ccgcaagcag  | aaccccgaca               | regegateta | ccagtacatg | gaccacctgt | acgtgggcag | 4200  |
| chacceggag  | atcggccagc               | accycaccaa | gategaggag | etgegeeage | acctgctgcg | 4260  |
| etggggette  | accacccccg               | acaagaagca | ccagaaggag | cccccttcc  | tgtggatggg | 4320  |
| ctacgagetg  | cacccgaca                | agrggaccgr | gcagcccate | gracracca  | agaaggacag | 4380  |
| cuggaccgug  | aacgacatcc               | agaagetggt | gggcaagetg | aactgggcca | gccagatcta | 4440  |
| egeeggeace  | aaggtgcgcc               | agetgtgcaa | getgetgege | ggcaccaagg | ccctgaccga | 4500  |
|             | ctgaccgagg               |            |            |            |            |       |
|             | cacggcgtgt               |            |            |            |            |       |
|             | ggccagtgga               |            |            |            |            |       |
| eggeaagtac  | gcccgcatga               | agggcgccca | caccaacgac | grgaagcagc | tgaccgaggc | 4740  |
| cgtgtagaag  | atcgccaccg               | agageategt | gatetgggge | aagaccccca | agttcaagct | 4800  |
| geecaceeag  | aaggagacct               | gggaggcctg | grggaccgag | tactggcagg | ccacctggat | 4860  |
| ccccgagtgg  | gagttcgtga               | acacccccc  | cctggtgaag | ctgtggtacc | agctggagaa | 4920  |
| ggageceate  | atcggcgccg               | agacetteta | cgcggacggc | gccgccaacc | gcgagaccaa | 4980  |
| gctgggcaag  | gccggctacg               | rgaccgaccg | cggccgccag | aaggtggtgc | ccctgaccga | 5040  |
| caccaccaac  | cagaagaccg               | agctgcaggc | catccacctg | gccctgcagg | acagcggcct | 5100  |
| ggaggtgaac  | atcgtgaccg               | acagccagta | caccctaaac | atcatccagg | cccagcccga | 5160  |
| caagagcgag  | agcgagctgg               | tgagccagat | catcgagcag | ctgatcaaga | aggagaaggt | 5220  |
| gtacctggcc  | tgggtgcccg               | cccacaaggg | catcggcggc | aacgagcagg | tggacggcct | 5280  |
| ggtgagcgcc  | ggcatccgca               | aggtgctgtt | cctggacggc | atcgacaagg | cccaggagga | 5340  |
| gcacgagaag  | taccacagca               | actggcgcgc | catggccagc | gacttcaacc | tgccccccgt | 5400  |

| aataaaaaaa  | gagatcgtgg | coanctacas | caantaccaa | ctgaagggg  | aggccatgca | 5460   |
|-------------|------------|------------|------------|------------|------------|--------|
| ggtggttaag  | gactgcagcc | ccagccgcga | caagegeeag | togaagggeg | tagaaaaaaa | 5520   |
| cggccaggcg  | gactgcagcc | coggeaceeg | coastacate | agaaccaaca | tastcccac  | 5580   |
| ggtgattettg | gtggccgtgc | acgragecas | agtassasta | gaggeegagg | agaccatass | 5640   |
| cgagaccggc  | caggagaccg | cetaetteet | gergaagerg | geeggeegee | ggcccgcgaa | 5700   |
| gaccgcgcac  | accgacaacg | gcagcaactt | caccagcacc | accycyaayy | cegeeegeeg | 5760   |
| gtgggccggc  | atcaagcagg | agtteggeat | cccctacaac | eeeeagagee | agggcgcgac | 2/00   |
| cgagagcatg  | aacaaggagc | tgaagaagat | categgeeag | grgcgcgacc | aggccgagca | 5820   |
| cctgaagacc  | gccgtgcaga | tggccgtgtt | catccacaac | ttcaagcgca | agggcggcat | 5880   |
| cggcggctac  | agcgccggcg | agcgcatcgt | ggacatcatc | gccaccgaca | tccagaccaa | 5940   |
| ggagctgcag  | aagcagatca | ccaagatcca | gaacttccgc | gtgtactacc | gcgacagccg | 6000   |
| cgaccccgtg  | tggaagggcc | ccgccaagct | gctgtggaag | ggcgagggcg | ccgtggtgat | 6060   |
| ccaggacaac  | agcgacatca | aggtggtgcc | ccgccgcaag | gccaagatca | tccgcgacta | 6120   |
| cggcaagcag  | atggccggcg | acgactgcgt | ggccagccgc | caggacgagg | acatgggcag | 6180   |
| caagtggagc  | aagagcagca | tcgtgggctg | gagggaggtg | agggagaggc | tgaggcagac | 6240   |
| cctggccgcc  | gccagaggcg | tgggcgccgt | gagccaggac | ctggagaagt | acggcgccgt | 6300   |
| gaccagcagc  | aacgtgaacc | accctagctg | cgcctggctg | gaggcccagg | aggaggagga | 6360   |
| ggtgggcttc  | cctgtgaggc | ctcacgtgcc | tctgaggcct | atgacctaca | agggcgccct | 6420   |
| ggacctgagc  | cacttcctga | aggagaaggg | cggcctggac | ggcctgatct | acagcaggag | 6480   |
| gaggcaggag  | atcctggacc | tgtgggtgta | caacacccag | ggctacttcc | ctgactggca | 6540   |
| gaactacaca  | cccggccctg | gcgtgaggta | ccctctgacc | ttcggctggt | gcttcaagct | 6600   |
| ggtgcctgtg  | gaccctgacg | aggtggagca | ggccaacgag | ggcgagaaca | acagcctgct | 6660   |
| gcaccctatg  | tgccagcacg | gcatggacga | cgaggagaag | gagaccctga | ggtggaggtt | 6720   |
| cqacaqcaqq  | ctggccctga | ggcacagggc | ccaggagatg | caccctgagt | tctacaagga | 6780   |
| ctactaaaca  | cgtggcgaat | tcttqctqtq | ccttctagtt | gccagccatc | tgttgtttgc | 6840 . |
| ccctccccq   | tgccttcctt | gaccctggaa | ggtgccactc | ccactgtcct | ttcctaataa | 6900   |
| aatqaqqaaa  | ttgcatcgca | ttqtctqagt | aggtgtcatt | ctattctggg | gggtggggtg | 6960   |
| gggcagcaca  | gcaaggggga | ggattgggaa | gacaatagca | ggcatgctgg | ggatgcggtg | 7020   |
| ggctctatgg  | gtacccaggt | gctgaagaat | tgacccggtt | cctcctgggc | cagaaagaag | 7080   |
| caggicacate | cccttctctg | tgacacaccc | tgtccacgcc | cctggttctt | agttccagcc | 7140   |
| ccactcatag  | gacactcata | gctcaggagg | gctccgcctt | caatcccacc | cgctaaagta | 7200   |
|             | tetetecete |            |            |            |            |        |
| qqaaqaaatt  | aaagcaagat | aggctattaa | gtgcagaggg | agagaaaatg | cctccaacat | 7320   |
| gtgaggaagt  | aatgagagaa | atcatagaat | ttetteeget | tectegetea | ctgactcgct | 7380   |
| acactcaatc  | gttcggctgc | ggcgagcggt | atcagctcac | tcaaaggcgg | taatacggtt | 7440   |
| atccacagaa  | tcaggggata | acqcaqqaaa | gaacatgtga | gcaaaaggcc | agcaaaaggc | 7500   |
| caggaaccgt  | aaaaaggccg | cattactaac | gtttttccat | aggeteegee | cccctgacga | 7560   |
| gcatcacaaa  | aatcgacgct | caaqtcaqaq | gtggcgaaac | ccgacaggac | tataaagata | 7620   |
| ccaggcgttt  | cccctggaa  | acteceteat | gcgctctcct | gttccgaccc | tgccgcttac | 7680   |
| cogatacctg  | teegeettte | tecetteggg | aagcgtggcg | ctttctcaat | gctcacgctg | 7740   |
| taggtatete  | agticggtgt | aggtcgttcg | ctccaaqctq | ggctgtgtgc | acgaaccccc | 7800   |
| cottcaoccc  | gaccgctgcg | ccttatccqq | taactatcqt | cttgagtcca | acccggtaag | 7860   |
| acacgactta  | tegecaetgg | caqcaqccac | tggtaacagg | attagcagag | cgaggtatgt | 7920   |
| aggcggtgct  | acagagttct | tgaagtggtg | gcctaactac | qqctacacta | qaaqqacaqt | 7980   |
|             | tgcgctctgc |            |            |            |            |        |
| atccggcaaa  | caaaccaccg | ctaataacaa | taatttttt  | gtttgcaagc | agcagattac | 8100   |
| gcgcagaaaa  | aaaggatctc | aagaagatcc | tttgatcttt | tetacqqqqt | ctgacgetea | 8160   |
|             | aactcacgtt |            |            |            |            |        |
| ctagatectt  | ttaaattaaa | aatgaagttt | taaatcaatc | taaagtatat | atgagtaaac | 8280   |
| ttaatetaae  | agttaccaat | gcttaatcag | tgaggcacct | atctcagcga | tctgtctatt | 8340   |
| teathcatee  | atagttgcct | gactccgggg | adadaaaaca | ctgaggtctg | cctcataaaa | 8400   |
| aaggtottoc  | tgactcatac | caggeetgaa | tegecocate | atccadccad | aaagtgaggg | 8460   |
|             | gatgagagct |            |            |            |            |        |
|             | acggtctgcg |            |            |            |            |        |
|             | tattcaacaa |            |            |            |            |        |
| ttacaaccaa  | ttaaccaatt | ctgattagaa | aaactcatco | agcatcaaat | qaaactqcaa | 8700   |
|             |            | 2          |            | -          |            |        |

| tttattcata | tcaggattat | caataccata | tttttgaaaa | agccgtttct | gtaatgaagg | 8760  |
|------------|------------|------------|------------|------------|------------|-------|
| agaaaactca | ccgaggcagt | tccataggat | ggcaagatcc | tggtatcggt | ctgcgattcc | 8820  |
| gactcgtcca | acatcaatac | aacctattaa | tttcccctcg | tcaaaaataa | ggttatcaag | 8880  |
| tgagaaatca | ccatgagtga | cgactgaatc | cggtgagaat | ggcaaaagct | tatgcatttc | 8940  |
| tttccagact | tgttcaacag | gccagccatt | acgctcgtca | tcaaaatcac | tcgcatcaac | 9000  |
| caaaccgtta | ttcattcgtg | attgcgcctg | agcgagacga | aatacgcgat | cgctgttaaa | 9060  |
| aggacaatta | caaacaggaa | tcgaatgcaa | ccggcgcagg | aacactgcca | gcgcatcaac | 9120  |
| aatattttca | cctgaatcag | gatattcttc | taatacctgg | aatgctgttt | tcccggggat | 9180  |
| cgcagtggtg | agtaaccatg | catcatcagg | agtacggata | aaatgcttga | tggtcggaag | 9240  |
| aggcataaat | tccgtcagcc | agtttagtct | gaccatctca | tctgtaacat | cattggcaac | 9300  |
| gctacctttg | ccatgtttca | gaaacaactc | tggcgcatcg | ggcttcccat | acaatcgata | 9360  |
| gattgtcgca | cctgattgcc | cgacattatc | gcgagcccat | ttatacccat | ataaatcagc | .9420 |
| atccatgttg | gaatttaatc | gcggcctcga | gcaagacgtt | tcccgttgaa | tatggctcat | 9480  |
| aacacccctt | gtattactgt | ttatgtaagc | agacagtttt | attgttcatg | atgatatatt | 9540  |
| tttatcttgt | gcaatgtaac | atcagagatt | ttgagacaca | acgtggcttt | cccccccc   | 9600  |
| ccattattga | agcatttatc | agggttattg | tctcatgagc | ggatacatat | ttgaatgtat | 9660  |
| ttagaaaaat | aaacaaatag | gggttccgcg | cacatttccc | cgaaaagtgc | cacctgacgt | 9720  |
| ctaagaaacc | attattatca | tgacatťaac | ctataaaaat | aggcgtatca | cgaggccctt | 9780  |
| tcgtc      |            |            |            |            |            | 9785  |